



CLASIFICADOR DE EMPRESAS BAJO CRITERIOS DE LA EBC

Estudio de Aplicación

Descripción breve

Estudio sobre la aplicación de tecnologías de aprendizaje automático en la valoración de empresas que realiza actualmente, por medio de auditorías, la organización de la Economía del Bien Común

Alejandro Layos Montero
Tutores: David Palomar y Miguel Ángel Patricio

Resumen

Mi trabajo se centra en la aplicación de nuevas tecnologías al proceso de clasificación desarrollada por la teoría empresarial de la Economía del bien común. Esta teoría económica desarrollada en Austria ha ido creando un proceso de clasificación empresarial alternativo al actual. En este proceso se da un peso primordial a criterios éticos y ecológicos frente a los económicos. Durante años han desarrollado un sistema de auditorías de mano de personas expertas. El objetivo es poder lograr un modelo de desarrollo económico en el que las empresas estén interconectadas y se beneficien unas y otras y a su vez aporten un producto a los consumidores que ayude en la consecución del bien común.

En este proceso, que actualmente es manual, he visto una oportunidad de aplicar alguna tecnología que sirva para ahorrar tiempo a priori y adelantar trabajo previo. Las valoraciones de cada empresa se llevan a cabo bajo criterios más o menos normalizados pero sujetos a la interpretación de cada persona que audita. De esta forma se establecen distintas puntuaciones a las empresas en torno a 17 puntos concretos. Cada uno de estos puntos lleva una guía asociada con distintos criterios que han de tener más o menos peso en la valoración. Es en este momento donde entra la valoración de la auditora y es aquí donde he visto una oportunidad de estudiar la aplicación de una tecnología de aprendizaje automático.

La idea principal es estudiar la posibilidad de automatizar estas valoraciones a través de algoritmos de predicción. Analizando todos los documentos de valoración previos se pueden obtener numerosos indicadores y atributos sobre que trabajar para intentar predecir las valoraciones finales. Con estos atributos se realiza una minería de datos en torno a la información aportada por las empresas. Junto a esto y la utilización de algoritmos de aprendizaje automático se busca crear un modelo de predicción, previo a la auditoria, que sirva de borrador inicial e incluso, en un futuro y con más datos, como sistema de valoración de empresas sin la necesidad de horas de análisis.

Los resultados de esta investigación servirán a la organización como primeros pasos en este tipo de aplicación.

Contenido

1.Introducción y Objetivos	5
1.1. Descripción del proyecto	5
1.2. Objetivos Personales.....	6
1.3. Fases de Desarrollo	7
1.4. Medios Empleados.....	8
1.5. Estructura de la Memoria	8
2.Estados del Arte.....	9
2.1. ¿Qué es la economía del Bien Común?	9
2.2. Balance del bien común.....	12
2.2.1. Estructura	13
2.2.2. Funcionamiento	13
2.3. Matriz del Bien Común	16
2.3.1. A1 Gestión ética de los suministros.	17
2.3.2. C1 Calidad del puesto de trabajo e igualdad	18
2.3.2. C3 Promoción del Comportamiento Ecológico de las Personas Empleadas. 20	
2.4. Machine Learnig.....	21
3.Propuesta de Estudio.....	23
3.1. Planteamiento del Problema	23
3.2. Propuesta de Solución	24
4. Investigación.....	25
4.2. Modelación de datos	25
4.1.1. Modelo A1 Gestión ética de los suministros.....	26
4.1.2. Modelo C1 Calidad del Puesto de Trabajo e Igualdad	30
4.1.3. Modelo C3 Promoción del comportamiento ecológico de las personas empleadas	32
4.2. Algoritmos investigación.....	34
4.2.1. Funciones	35
4.2.2. Árboles de decisión	38
4.3. Entorno de Desarrollo.....	39
4.3.1. Hardware.....	39
4.3.2. Software	39
4.4. Pruebas a realizadas	40
4.4.1 Conjuntos de Datos A1.....	40

4.4.2. Pruebas A1	42
4.4.3. Conjuntos de Datos C1	43
4.4.4. Pruebas C1.....	44
4.4.5. Conjuntos de Datos C3	44
5. Resultados	46
5.1. A1	46
5.2. C1	50
5.3. C3	51
6. Análisis del sistema.....	54
6.1. Casos de uso.....	54
6.2. Requisitos.....	57
6.2.1. Requisitos Funcionales	58
6.2.2. Requisitos No funcionales	58
7. Diseño la Demo.....	61
7.1. Entorno de Desarrollo.....	61
7.2. Clasificador.....	61
7.2.1. Interfaz	63
7.2.2. Evaluador.....	64
8. Plan Económico	65
8.1. Planificación del Proyecto.....	65
8.2. Diagrama de Gantt.....	67
8.3. Costes.....	68
8.3.1. Personal.....	68
8.3.2. Hardware.....	68
8.3.3. Software	69
8.3.4. Total.....	69
8.4 Presupuesto Detallado.....	70
9. Conclusiones.....	71
10. Trabajo futuro.....	74
11. Bibliografía.....	75
Anexo I. Datos A1	76
Anexo II. Datos C1.....	81
Anexo II. Datos C3.....	85

1. Introduction and objectives	90
1.1. Description	90
1.2. Personal objectives	91
1.3. Phases of development.....	92
1.4. Media	93
2. Common Good economical theory.	93
2.1. Introduction.	93
2.2. Common Good Balance.	93
2.3. Common Good Matrix	94
3.Study proposal.....	95
3.1. Problem Statement.....	95
3.2. Solution proposal	96
4.Investigation	97
5.1Results	97

1.Introducción y Objetivos

En este primer punto intentare dar una visión general del trabajo que he realizado y que recojo a lo largo del documento. La idea es plantear una primera descripción que sirva de introducción y sitúe en el contexto a la persona que lo esté leyendo.

1.1. Descripción del proyecto

En primer lugar, lo más importante es definir en qué consiste mi proyecto, al menos en líneas generales. Como queda ilustrado en el título del documento, se trata desarrollar un sistema capaz de clasificar las empresas en torno a los criterios de la teoría económica del bien común. Este trabajo de estudio y de desarrollo se realizará mediante técnicas de minería de datos y aprendizaje automático. Esta teoría aporta un modelo de valoración de empresa que se aparta de la norma y plantea una clasificación basada en el bien social y la sostenibilidad. La teoría aporta herramientas para poder clasificar a las empresas que son utilizadas por las auditoras para valorarlas de forma adecuada.

La organización de la EBC tiene elaborados documentos y realiza formaciones a las personas interesadas para auditar empresas. A través de estas personas se dedican a realizar auditorías sociales a las empresas que se presten a ello. Cada uno de los indicadores tiene unas reglas y varemos asociado que sirve para dar un valor más o menos preciso. Este proceso de clasificación es costoso y lento y no siempre se dispone de personas suficientes para desplazarse a las acompañar y elaborar los análisis. Un problema añadido es que cada proceso de valoración es particular y depende de distintos factores, entre ellos de la persona que se encargar de evaluar. Pese a que tienen unas pautas para valorar cada uno de los indicadores, estos están sujetos a interpretación y no son del todo exactos.

La idea y el objetivo de este proyecto es investigar si se puede crear un sistema de valoración automático que permita, al menos, realizar valoraciones previas. Puesto que se dispone de evaluaciones previas y resultados, se decide trabajar a partir de estos para obtener un modelo de predicción que se acerque a la realidad. El sistema teorizado es aplicado por estas personas y no resulta ser tan preciso en la práctica, por lo tanto, se decide trabajar a partir de los ejemplos y no del marco teórico.

Es importante destacar el sentido del trabajo. No se busca obtener un modelo final con todos los indicadores. Se trata de investigar a partir de unos pocos indicadores si tiene sentido buscar la automatización del proceso.

Una de las partes importantes y previa en este proyecto es la de estudiar todos los documentos de evaluación e intentar modelar un conjunto de datos que sirvan de estudio y que, más adelante, puedan aportar conclusiones para el sistema. El proceso de estudio de los documentos es manual y consiste en leer y sacar todas las medidas o iniciativas que se reflejan y que afectan a los indicadores.

El objetivo final de este proyecto es aplicar diversas técnicas y algoritmos de machine learning a todos los datos previamente recogidos para intentar obtener un sistema automático de clasificación. Por ahora nadie se había interesado en aplicar estas técnicas al servicio de esta teoría y por lo tanto supone un primer paso en el camino a la automatización completa.

1.2. Objetivos Personales

Estos son los objetivos técnicos que busco en la realización del proyecto. A través de ellos quiero iniciar un proceso de aprendizaje del que me quiero aprovechar de cara al mundo laboral y a mi carrera profesional

- Aplicación de algoritmos de Machine learning al proceso de evaluación.
- Prototipado de un evaluador automático.
- Automatización del sistema de evaluación de la EBC.

Explicaré, también, las motivaciones personales que me han llevado hasta aquí puesto que sirven para comprender el sentido del trabajo.

- Profundización en técnicas de minería de datos
- Profundización en técnicas de machine learning
- Aprendizaje de trato con clientes
- Gestión de un proyecto en campo
- Aprendizaje de gestión de un proyecto de investigación

La economía del bien común plantea una alternativa a la valoración actual que se hace de las empresas. La realidad es que las empresas se valoran y se escogen en función de los ingresos y su rentabilidad. En esta lógica de mercado, generalmente son las personas que tiene acciones las que se benefician principalmente de las acciones de estas empresas. En esta tesitura se toman decisiones que muchas veces van en contra de las personas que están empleadas en la empresa. Y más allá de esto, se toman decisiones que van en contra del planeta y la sostenibilidad a largo plazo.

Planteando esta premisa, creo que es necesario otra forma de valorar y de dar ventajas a las empresas. Aquí es donde juega un papel fundamental la teoría del bien común. Pese a que no deja de ser una solución intermedia y reformista del sistema de producción actual nos aporta un punto de vista distinto. En el capítulo siguiente entraré un poco más en materia, pero consideraba importante dar unas pinceladas para mostrar mis motivaciones para tratar este tema.

Claramente esta alternativa necesita inversión y dedicación. Con este proyecto busco aportar mi grano de arena en esta corriente y proveerles de sistema de clasificación que automatice su teoría. De esta forma la gente podrá acceder con facilidad y servirse de estas clasificaciones para consumir de forma consciente y sostenible.

1.3. Fases de Desarrollo

1. Estudio: el punto de comienzo se basa en llevar a cabo un proceso de estudio y de indagación.

En primer lugar, había que investigar todo lo referente a la economía del bien común. Era algo que conocía, pero no en profundidad y tuve que ponerme a leer e investigar su historia. Durante este proceso mantuve contacto con la organización en Madrid y me reuní varias veces con las personas que la integran. A través de estas reuniones conocí un poco mejor su labor y me hice una idea más precisa de la teoría que esgrimen. A través de reuniones y de estudiar los documentos y videos que tiene colgados obtuve el conocimiento necesario.

En segundo lugar, quería refrescar y profundizar un poco más en mis conocimientos sobre aprendizaje automático para afrontar el desarrollo del proyecto de forma adecuada. Repasé las asignaturas que había cursado sobre este tipo de tecnología (aprendizaje automático e inteligencia artificial). Para ampliar un poco más realice un curso de machine learning a través de la plataforma coursera. Como se puede observar este proceso ha sido de los más largos junto a la investigación.

2. Reconocimiento del problema. Una vez adquiridos los conocimientos necesarios para afrontar mi objetivo me dediqué a encontrar de qué forma o como podría aplicar esta tecnología al proceso de evaluación. Revisando toda la documentación y leyendo los documentos de las empresas, me di cuenta de dos cosas que afectaban a estas. La primera el tiempo que requiere una evaluación completa de cada empresa. En algunos casos las auditorías duraban 100 horas. El segundo problema resulta de la propia evaluación. Observando las auditorías se observa que, aunque los criterios están preestablecidos no son del todo precisos y no se aplican de la misma forma en las empresas. Esta fase no me llevo demasiado tiempo, debido a que me fui dando cuenta a la vez que investigaba sobre el tema en cuestión.
3. Análisis. Una vez reconocido el problema me dedique a recabar datos a través de los documentos aportados por la EBC sobre las empresas. A través de una lectura exhaustiva de los mismo fui obteniendo destinos datos o atributos que estuvieran (o que no) las auditorías y que a priori hubieran servido para la valoración. Esta fase ha sido lenta y laboriosa puesto que tuve que leer detenidamente e ir sacando conclusiones de todos los PDF que me había facilitado. Poco a poco fui sacando suficiente información para pasar a la fase de diseño de los experimentos.
4. Diseño. Esta fase consistió básicamente en elaborar un plan de pruebas y en generar el modelo sobre el que iba a trabajar. En cuanto a las pruebas

básicamente designe que algoritmos de machine learning iba a probar y que tipo de pruebas iba a hacer con ellos. En el diseño del modelo me centre en definir bien los atributos que había recogido en la fase anterior y elaborar un modelo en entradas y salida acorde a lo que necesitaba. Este modelo sufrirá modificaciones a lo largo de la investigación, pero me servía de prototipo en primera instancia.

5. Investigación Una vez dejada todo definido me dediqué a investigar con los datos de los que disponía a través de la plataforma weka. Fui realizando las pruebas que me programé en un principio y luego añadiendo alguna más que me resultaba útil sobre la marcha. A través de esta investigación obtuve las conclusiones pertinentes para desarrollar el sistema de predicción.
6. Desarrollo. Esta última parte consistió básicamente en desarrollar una pequeña aplicación para mostrar cómo funciona la ecuación de predicción. Esta parte es básicamente visual, la importancia radica en la investigación y en la búsqueda de una fórmula de predicción.
7. Redacción. Esta fase ha sido transversal y he ido redactando la memoria conforme trabajaba en el resto de tareas.

1.4. Medios Empleados

- Equipo Físico
 - Ordenador Sobremesa: Procesador AMD FX™-8320 Ocho núcleos 3.5 GHz, RAM 8GB, Disco Duro 1 T B y Disco SSD 256 GB.
 - Ordenador Portátil: Asus, Procesador Intel® Core™ i5-6300HQ Quad-Core (6M Cache, 2.3GHz hasta 3.2GHz), RAM 8 GB y Disco Duro 1 TB
- Software
 - Oficce 2016
 - Weka 3.8
 - Windows 10
 - Notepad++

1.5. Estructura de la Memoria

En este punto haré un pequeño resumen de la estructura que voy a seguir a lo largo de la memoria.

- **Estado del Arte.** En este primer punto se recoge el contexto en el que se encaja el proyecto. A lo largo de este punto queda reflejada y explicada a grandes rasgos la teoría económica del bien que sirve como punto de partida y de estudio en mi trabajo. También se expone el proceso de valoración de empresas, así como, la matriz que sirve para comparar unas empresas con otras y supone parte central

en esta teoría. Para terminar este punto hago un pequeño recorrido sobre que es Machine Learning y que aplicaciones tiene.

- **Propuesta de Estudio.** Una vez expuesto el contexto, durante este punto se expone el problema encontrado y el objetivo central de este trabajo. Se explica el problema concreto y que ventajas tendría solucionarlo. Luego un desarrollo un poco más técnico de que solución informática se ha investigado y que implicaciones tiene.
- **Pruebas/Investigación.** En el plan de pruebas se refleja todo el proceso de investigación que se ha llevado a cabo. Se recogen todas las pruebas realizadas, los cambios en el modelo, los errores encontrados etc. Primero se expone la retahíla de algoritmos que se han utilizado para investigar el modelo de predicción, cada uno contará con un pequeño desarrollo teórico. Después se da cuenta del entorno de desarrollo escogido. Se termina detallando el modelo, que será un prototipo, con el que se va a investigar. Este modelo constará de atributos y salidas nominales.
- **Resultados:** se exponen todos los datos obtenidos en la investigación y los sistemas de predicción encontrados. A partir de estos datos se desarrollará el consiguiente sistema.
- **Análisis del Sistema.** Aquí se estudian los casos de uso, requisitos y el diagrama de clases del que constará la aplicación a desarrollar.
- **Diseño del Sistema.** En este punto se recoge el desarrollo técnico del sistema a implementar
- **Plan Económico.** Aquí se expone una pequeña memoria económica.
- **Conclusiones.** Se muestran las conclusiones del estudio en el contexto actual y si tiene sentido su aplicación. También mi visión personal sobre este avance o que puede aportar a este proyecto y la experiencia en el desarrollo del proyecto.
- **Trabajo Futuro.** En este último punto se dan unas pinceladas de lo que puede ser el trabajo futuro en esta línea de trabajo. Si se decide continuar, da una pequeña estrategia o ayudas para seguir por donde se ha iniciado.
- **Bibliografía**

2.Estados del Arte

2.1. ¿Qué es la economía del Bien Común?

Comenzaré el capítulo con una pregunta ¿Qué es la economía del bien común? Si buscamos en internet, tras una breve consulta en internet encontramos la primera definición:

” ECONOMÍA DEL BIEN COMÚN ES UN PROYECTO ECONÓMICO ABIERTO A LAS EMPRESAS Y PROMOVIDO POR EL ECONOMISTA AUSTRÍACO CHRISTIAN FELBER QUE PRETENDE IMPLANTAR Y DESARROLLAR UNA VERDADERA ECONOMÍA SOSTENIBLE Y ALTERNATIVA A LOS MERCADOS FINANCIEROS EN LA QUE NECESARIAMENTE TIENEN QUE PARTICIPAR LAS EMPRESAS”

La teoría del bien común nos plantea un modelo económico alternativo al actual. Nos presenta un modelo que se centra en la sostenibilidad, la ecología, la ética empresarial y los buenos tratos y que se aparta de la concepción normalizada de la empresa. Esta teoría se inicia de manos del Austriaco Christian Felber ¹. El mismo la define como una alternativa al capitalismo de mercado y a la economía planificada. Esta teoría supone una revisión del modelo económico actual y le aporta una visión largoplacista en la que garantiza la sostenibilidad y el desarrollo del bien común para nuestro planeta.

A lo largo de este punto esbozare las características principales de esta teoría. No quisiera plantear el debate acerca la viabilidad de esta teoría o de las posibilidades reales de implantación en el momento actual y que inconvenientes podría tener. A lo largo del proyecto me ceñiré a realizar una pequeña aportación a esta corriente económica y las organizaciones y empresas que se adscriben a ella.

Entrando en materia, para explicar en qué consiste esta teoría tomaré prestada la explicación en 20 pasos que el propio Christian Felber aporta². (En su origen en enlace utilizado fue otro, pero desapareció y encontré otro más actualizado) Por lo tanto, ¿En qué consiste la economía del bien común?

1. Los valores en los que se fundamenta y a los que, por lo tanto, da prioridad son: la cooperación, el aprecio, la democracia y la solidaridad. Esto se antepone a los imperantes en el modelo actual.
2. La legalidad y el marco normativo se adaptan a esta nueva escala de valores. Ahora en vez de primar la competencia y las intenciones meramente lucrativas se beneficia la cooperación entre empresas.
3. La forma de medir el éxito y la rentabilidad de una empresa cambian sustancialmente. Antiguos indicadores como el TIR o el Beneficio pierden importancia a cambio de nuevos valores. Se generaliza el balance del bien común a través del cual se priorizan nuevos indicadores como la sostenibilidad, la ecología o la cooperación.
4. Toda esta inversión y preocupación por los nuevos valores tiene recompensa. Cuanto mejores valoraciones tengan y más esfuerzos dediquen a aplicar estos

¹ Wikipedia. Christian Felber[en línea] https://es.wikipedia.org/wiki/Christian_Felber

² EBC. La idea [en línea] <http://economia-del-bien-comun.org/es/content/contenido>

cambios a sus empresas a mejores excepciones fiscales aspirarán. De tal forma que si una empresa tiene mejores valoraciones podrá tener, por ejemplo, créditos a menores intereses. De esta forma se incentivan el cambio.

5. El beneficio financiero pasa automáticamente a ser un medio. Las empresas no se centran en el beneficio si no en los resultados que provocan con su actividad comercial. Usan ese beneficio para lograrlos. El beneficio, como tal, ha de ser reinvertido.
6. Puesto que ahora no hay necesidad de aumentar beneficios constantemente, tampoco hay necesidad de crecer por el hecho de hacerlo. Las empresas pueden mantener su tamaño de acuerdo a su actividad comercial sin miedo a ser adquiridas o fagocitadas por otras más grandes.
7. Debido a que no existe este afán de crecimiento constante, cada empresa puede dedicarse a su sector y existirán múltiples ramas. Estas empresas pueden cooperar entre ellas para obtener resultados conjuntos sin miedo de ser perjudicadas y con responsabilidad social.
8. Se establecen leyes específicas para evitar la concentración de riqueza. Los ingresos personales no pueden superar cierta cantidad. Esta cantidad se establece en función del salario mínimo. También se establecen límites en la propiedad individual. Las herencias y cesiones familiares están limitadas a una cantidad específica. Todo lo que exceda estos límites se destinará al llamado “fondo de generaciones” que se destinará a reducir la desigualdad estructural del sistema.
9. Cuando las empresas alcanzan cierto tamaño sus decisiones pasan parcialmente a la sociedad. Se podrán organizar “parlamentos regionales.
10. Esto mismo es aplicable a los bienes democráticos. Lo que ahora consideramos como servicios públicos pasaría a ocupar esta nueva categoría.
11. El banco central juega un papel importante en todo este entramado. Será el encargado de monopolizar la actividad monetaria. Se garantizan los préstamos a intereses reducidos, las ayudas para emprender, las cuentas corrientes... El estado no dispone de poder decisorio sobre este, se ampara en la soberanía popular. Además, el estado se nutre de créditos de este banco para llevar a cabo su labor.
12. Se establecen 3 niveles monetarios. Una moneda global para las transacciones internacionales. Una moneda nacional para las nacionales. Y distintas monedas regionales para las locales. De esta forma se evitan las desigualdades entre países y se protege en cierta medida el comercio local.
13. Desaparece la propiedad privada aplicada a la naturaleza. Por lo tanto, no se podrá especular con todo el suelo y tampoco se podrá vender. Toda persona que necesite de un espacio ya sea para un negocio o una vivienda se le cederá de forma gratuita o con una tasa, en función de sus necesidades.
14. Como fin último de la producción se establece la reducción de la huella ecológica. El crecimiento económico pasa a un plano muy secundario para garantizar un nivel de sostenibilidad y justicia global óptimos.

15. En materia laboral, se produce una reducción drástica de la jornada y las horas laborales exigidas. Tendería a 30 horas semanales. De esta forma se saca a la luz y se da importancia al trabajo de los cuidados, el personal, y el político o de actividades públicas. Con esta solución el trabajo cede espacio al desarrollo personal y comunitario.
16. Acorde a este tipo de filosofía cada 10 años se premia al trabajador con un año sabático íntegramente financiado. Esto ayuda a descargar el mercado laboral a su vez.
17. Se plantea un cambio de modelo de estado. Pasaríamos de democracias representativas a modelo mixtos donde se intercalarán con democracias directas o participativas. Obviamente en caso de no existir democracias también se extendería a este modelo. De esta forma se modificaría las constituciones para dotar de poder al poder constituyente más allá del expresado en parlamentos comunes. Este poder podrá verse reflejado en la toma de decisiones sobre los bienes públicos.
18. Se aboga por el consenso y la soberanía nacional para lograr un acuerdo en las leyes que afecten a la población. Es decir, cada una de las decisiones en materia legislativa han de ser refrendadas mediante consulta a todo el territorio que afecten. En este aspecto, se plantean distintas cámaras o asambleas competentes en materias específicas y no una general o dos como se estila en la actualidad.
19. Este cambio ha de afianzarse, por supuesto, a través de la educación. El modelo educativo ha de repensarse para enseñar a las nuevas generaciones en valores del bien común.
20. Los valores empresariales cambian y por lo tanto también lo hace el perfil de las demandas laborales.

A grandes rasgos este sería el modelo económico, pero también cultural y social, que proponen desde la EBC. En los siguientes puntos me centraré en exponer las herramientas en las que se basa para provocar este cambio y las técnicas que se emplean para valorar a las empresas.

2.2. Balance del bien común

El balance del bien común es la herramienta principal de esta teoría. A través de este balance se puede comprobar el trabajo de la empresa en torno al bien común. Esta herramienta mide cuanto sirve una compañía al bien común. Su objetivo es hacer de este tipo de balance el principal medio para valorar a las empresas y que sirva para comparar y como ayuda al consumidor a escoger, al estado a contratar y a otras empresas a asociarse. Tendría la misma fuerza que tiene el balance actual, solo que se centraría en otros puntos que ellos consideran decisivos en el largo plazo.

Como ya he explicado en el punto anterior este balance sustituye a los tradicionales económicos y rentabilidad. De hecho, es mucho más amplio puesto que mide muchos más factores que el puramente económico. La tesis principal es que el desempeño económico pasa a un segundo plano y no se considera un objetivo final. El dinero sería un medio para llevar a cabo un fin para la sociedad, aunque sea un producto de venta al uso. Con las ganancias económicas se consigue dotar de un servicio o un producto de utilidad, sostenible y ético. Por lo tanto, este balance viene a medir si se consiguen los objetivos a los que debería aspirar todas las empresas en el caso de que todas se midieran por este patrón.

Este balance elaborado en completo por esta organización. Es un proceso en constante revisión puesto que la realidad cambia y el mercado también. Por lo tanto, buscan constantemente como hacerlo más adecuado. En la actualidad este balance consta de 17 indicadores. Como exponen ellas mismas, no es proceso de evaluación precios, buscan que sea comprensible y que se pueda aplicar con facilidad.

2.2.1. Estructura

El balance consta de tres partes principales. Un informe, el propio balance y el testado. El informe es un documento extenso y propio de cada empresa en el que se detallan todas las actividades asociadas a cada uno de los 17 indicadores de la matriz. Nos presenta el estado general de la empresa y todos los esfuerzos que realiza para la consecución del bien común. También incluye fallos observados y propuestas de mejora, así como, intenciones de cambio a corto o medio plazo. Después tendríamos el balance en sí. Básicamente es el producto central del proceso y supone el instrumento principal de medición. Se trata de unas tablas de valoración para calcular las 17 puntuaciones finales de cada empresa. En el siguiente punto se ahondará mucho más en esta matriz, puesto que ella y el proceso de creación suponen la parte troncal de este proyecto. Por último, el testado sería el producto final. Es una matriz con los 17 indicadores. Es el elemento con más fuerza hacia fuera. Con esta matriz se puede comparar empresas, se pueden realizar históricos para comprobar la mejora, se puede observar a simple vista la calidad de una empresa...

2.2.2. Funcionamiento

Es importante detallar en que consiste este balance. Aunque en el siguiente punto lo hagamos pormenorizado, ahora se dará una explicación general.

El balance consta de 17 indicadores. Cada uno de estos indicadores tiene un valor de que oscila entre 20 y 90 puntos. Esta puntuación final se puede dar en puntos o en

porcentaje sobre el total. Para valorar cada uno de estos indicadores existen tablas internas con niveles que ayudan a aproximar los valores. Tablas como esta:

Categoría	Principiante (0-10%)	Avanzado (11-30%)	Experimentado (31-60%)	Ejemplar (61-100%)
Institucionalización	„Comienzo a pensar en ideas“	Primeras actividades serán aplicadas de forma puntual	Aspectos básicos serán dirigidos por medio de procesos institucionalizados	Estrategia y evaluación
Alcance en la estructura de la empresa	1 Emplazamiento	Para una parte de los trabajadores válido /planeada sucesiva ampliación	Válido para la gran mayoría de los trabajadores y de los emplazamientos (especialmente los de riesgo)	Válido para todos los trabajadores y emplazamientos
Alcance en productos/servicios	Sólo corresponde a pocos P/S (< 10%)	Corresponde a una parte significativa de los P/S (> 10%)	Corresponde a la gran mayoría de los P / S (> 50%)	Corresponde a toda la gama de P / S (> 90%)
Avance relativo	Ligeramente por encima de la media del sector	Ligeramente sobre la media del sector + se reconoce una tendencia positiva	Claramente sobre la media del sector+ tendencia claramente positiva	Liderazgo en el sector, si es definible, EBC-objetivo alcanzado
Cultura de empresa	Las primeras reuniones temáticas en la directiva , las primeras medidas de concienciación	Un comportamiento determinado será sólo promovido, porque/cuando no aparezca ninguna desventaja para la empresa	Esfuerzos amplios y objetivos definidos, la directiva vive los valores	Integrado, la directiva vive los valores, trato regular/ lematización del comportamiento

Estos porcentajes son niveles que van adquiriendo y cada una de las categorías tiene un peso específico en el indicador. De tal forma que al final se obtiene un porcentaje total con el que se calculan los puntos que van del 30 al 90. Cada empresa tiene un máximo de 1000. Hay también criterios negativos que ayudan a evitar las malas prácticas.

Estas tablas internas sirven como punto de apoyo en la valoración, pero como se puede observar no son precisas y usan elementos subjetivos. Cada uno de estos balances ha de ser trabajado y valorado, no siempre se siguen los mismos criterios. Una vez valorada la empresa se obtendría la matriz que veremos más adelante con cada uno de los indicadores. El funcionamiento como se puede apreciar es simple, puesto que solo se trata de evaluar en torno a 17 criterios, pero a la vez tiene varios grados de dificultad añadida como son la subjetividad y la amplitud de factores que influyen.


Siguiendo con el funcionamiento de este balance llegaríamos a la parte de la puntuación. ¿En qué influye tener mejor o peor puntuación y que implica? Bueno otras de las ideas de esta teoría es ser capaces de institucionalizarla e implicar a las administraciones. Del mismo modo que el estado prima a las empresas por determinadas actuaciones, como pueden ser contratar a parados de larga duración o personas con movilidad reducida, también tendrían ventajas fiscales cuanto más nota saquen. Estas fiscales irían desde la rebaja del IVA hasta la rebaja de impuestos más generales. Este incentivo valdría para motivar aún más a que realicen este tipo de balance. Esta puntuación también supondría un medio para que los consumidores

puedan comparar a la hora de escoger que producto comprar. Puede observar la tabla y elegir a que criterios quieren dar importancia a la hora de comprar. Otro detalle importante que supone esta valoración es que entre empresas también puede elegir con cuales colaborar y si encajan o no en sus criterios.

La ventaja de este balance es que se puede aplicar a cualquier empresa y solo existe una versión final. De esta forma se estandariza para todas.

Como ya se ha comentado con anterioridad. El balance se lleva a cabo como las actuales auditorías externas. En primer lugar, se realiza un balance interno de elaboración propia, después se lleva a las auditoras de la EBC que comprueban y estudian cada caso. Por último, elaborar el testado final y dan validez al balance.

2.3. Matriz del Bien Común

Valor ► Grupo de contacto ▼	Dignidad humana	Solidaridad	Sostenibilidad ecológica	Justicia social	Participación democrática y transparencia
A) Proveedores	 A1: Gestión ética de la oferta/suministros				90
B) Financiadores	B1: Gestión ética de finanzas				30
C) Empleados inclusive propietarios	C1: Calidad del puesto de trabajo e igualdad 90	C2: Reparto justo del volumen de trabajo 50	C3: Promoción del comportamiento ecológico de las personas empleadas 30	C4: Reparto justo de la renta 60	C5: Democracia interna y transparencia 90
D) Clientes / productos / servicios / otras empresas	D1: Relaciones éticas con los clientes 50	D2: Solidaridad con otras empresas 70	D3: Concepción ecológica de productos y servicios 90	D4: Concepción social de productos y servicios 30	D5: Aumento de los estándares sociales y ecológicos sectoriales 30
E) Ámbito social: región, soberanía, generaciones futuras, personas y naturaleza mundial	E1: Efecto social/ Significado del producto/ servicio 90	E2: Aportación a la comunidad 40	E3: Reducción de efectos ecológicos 70	E4: Orientación de los beneficios al bien común 60	E5: Transparencia social y participación en la toma de decisiones 30
Criterios – Negativos	Quebrantamiento de las Normas de trabajo OIT -200	Compra hostil -200	Gran impacto medioambiental a ecosistemas -200	Evasión de impuestos -200	No revelación de todas las participaciones -100

Como se puede apreciar tiene 17 indicadores cada uno centrado en un ámbito diferente. Esta sería la matriz final que ayuda a comprar a empresas. No voy a definir los 17 indicadores puesto que para eso ya tiene su propio manual. Pero me parece importante que defina los que voy a estudiar en este proyecto.

2.3.1. A1 Gestión ética de los suministros.

Este indicador trata de medir como se relacionan con los proveedores de su entorno y como gestionan sus propias compras. Está enfocado a que cada eslabón de la cadena de creación de valor conozca y elija a los que tiene inmediatamente al lado. De esta forma se asegura un trato familiar, ético, transparente y que repercute en el bien común. Su objetivo es que todas las empresas sean partícipes de la cadena de producción y elijan con criterios éticos aquellas empresas que lo merezcan. Es la mejor forma de asegurar que el producto final cumple con los estándares éticos o ecológicos necesarios. Si todas se implican el bien común está garantizado.

Su tabla de valoración interna sería esta:

Criterios	Principiante (10%)	Avanzado (30%)	Experimentado (60%)	Ejemplar (100%)
Consideración de aspectos regionales, ecológicos y sociales, alternativas de más valía (45%-65%)	Puntual para productos clásicos de riesgo (Eco-electricidad)	En algunos productos y servicios básicos	En la gran mayoría de productos y servicios básicos +reducción clara de sustancias críticas sin alternativa de más valía	En todos los productos y servicios básicos adquiridos +soluciones innovadoras para la reducción de sustancias críticas sin alternativa de más valía
Consideración activa de los riesgos de productos y servicios adquiridos y procesos de protección (25%-45%)	Consideración interna por medio de la recogida activa de informaciones sobre el tema Integración de aspectos sociales y ecológicos en los contratos/reglamentos (código de conducta/código ético)	Auditoría interna en caso de riesgos y de proveedores importantes Formaciones (seminarios, talleres, presupuesto para charlas de expertos) todos los trabajadores de proceso de ventas	Evaluación regular de riesgos y alternativas Garantía mediante auditoría independiente (p. ej. productos y servicios certificados bajo distintivos sociales o ecológicos, cooperación con ONGs)	Cooperación con proveedores y coempresarios respecto a aspectos sociales y ecológicos
Marco estructural para un precio justo (10%)	Renuncia a procesos de adquisición derivados de los precios (p. ej. subastas, licitaciones públicas) Ningún sistema de bonificación para compradores dependiente del precio de compra	Largas relaciones cooperativas, se priorizarán frente a cambiantes orientadas a los costes	Estructuras de ventas establecidas y directas Evaluación del comportamiento del vendedor por medio de conversaciones regulares, con el foco a los desafíos, derivados de la adquisición ética	Estructuras innovadoras en la adquisición (p. ej. participación en monedas o conceptos monetarios alternativos, etc.)

Una de las propuestas de evaluación que se da para este punto es la de realizar un estudio pormenorizado de todos los proveedores con pesos en función de la importancia. Como ya veremos más adelante pocas empresas lo llevan a cabo y será objeto de problemas en el proyecto

2.3.2. C1 Calidad del puesto de trabajo e igualdad

Este indicador trata de valorar la naturaleza de los puestos de trabajo. Su objetivo es que cada puesto de trabajo pueda asegurar que las personas puedan estar de acuerdo a sus posibilidades, ayudar a su desarrollo y generar un ambiente de comunidad dentro de la empresa. Se valora la autorganización, la conciliación familiar, la salud física, y otros indicadores que se centran en valorar las condiciones humanas del puesto de trabajo. Su tabla de valoración es esta:

Criterios	Principiante	Avanzado	Experimentado	Ejemplar
Horario de trabajo (20%) (Voluntario y decidido propiamente, no obligado por la empresa)	Horario de trabajo flexible y modelo de tiempo parcial	Horario de trabajo flexible y modelo de tiempo parcial + con la implicación del empleado	+ Descarga activa temporal de la ocupación de los niños (p. ej. guarderías, cuidadores/as)	Organización completamente propia del horario de trabajo (reparto del trabajo en función de los acuerdos de resultados)
Configuración del puesto de trabajo (10%) (para EUP 20%)	Todos los puestos de trabajo son ergonómicos (p. ej. luz y clima, etc.)	+ apto para minusválidos (acceso libre de barreras para todas las espacios); + Posibilidades extras para la elección libre del propio lugar de trabajo (p. ej. Homeoffice) + espacios adicionales para relajación y movimiento (p. ej. zonas verdes, cuarto de descanso)		
Salud física y seguridad (10%) (para EUP 20%)	Programas de prevención de salud y medidas de sensibilización (talleres y ponencias; 2 días p.a.)	+ (comienzo) promoción activa de la alimentación sana; (p. ej. ninguna máquina con "fast food", sino cestas biológicas)	+ ofertas individuales de asesoría en salud, chequeos voluntarios y acciones de deporte	+ (experimentado) Cocina Ecológica y cuidado (ev. por medio de autoabastecimiento o cocinar con profesionales)
Salud psíquica (15%) (para EUP 30%) número de días = por empleado, por año. Contenido en parte cambiable	Un día: Eventos de trabajadores para desarrollo de las capacidades de equipo y trabajo en parte de temas propios (p. ej. supervisión, coaching)	Dos días: talleres y formaciones regulares para competencia social (p. ej. Gestión de conflictos, comunicación sin violencia) y salud (reducción de stress)	Tres días: Talleres para formación continua (competencias sociales, desarrollo de la personalidad y mediadores)	Cuatro días: formación continua para el desarrollo personal y de competencias sociales
Organización propia, satisfacción en el puesto de trabajo, sentido (15%) (para EUP 30%)	Codeterminación y relación en la creación de las tareas (en equipo)	+ Encuesta de trabajadores (1x p.a.) sobre el puesto de trabajo y la cultura de confianza +medidas para fomentar la desjerarquización (2-3	Codeterminación en la creación de tareas, dirección y salarios, rotaciones del trabajo y enriquecimiento del campo de tareas gracias a la	Organización propia, cocreación de visión y estrategia de empresa, objetivo de reparto del trabajo (valor orientado a 1 nivel jerárquico)

		niveles jerárquicos)	formación específica	
Igualdad y trato igualitario para hombre y mujer (20%)	Género y Formaciones en diversidad, instalación de una persona responsable para los temas de género (a partir de 5 trabajadores) oferta de trabajo libre de discriminación, diversidad en el puesto de trabajo.	<ul style="list-style-type: none"> + Elaboración de un informe de igualdad, incluido el fomento de mujeres y de carencia paternal + (F) presupuesto de gender (incl. Transparencia de todos los salarios y de co-decisión de los trabajadores en el sueldo) encargado de igualdad de trato (1 pers. cada 25 empleados) + 50 % proporción de mujeres en la dirección de la empresa; formación y desarrollo continuo de personas de dirección (2 días p.a./p.p.) 		
Desfavorecidos (p. ej. personas con discapacitaciones, emigrantes, desempleados de largo tiempo) (10%)	Formaciones obligatorias sobre el tema anti-discriminación, búsqueda de personal adecuada (diversidad); cumplimiento en parte de la cuota legal (sobre 2%)	<ul style="list-style-type: none"> + Incorporación de personas o cooperaciones para el apoyo; cumplimiento de la cuota legal (=ningún pago compensatorio) + Formaciones en el trato de necesidades especiales; sobrecumplimiento de la cuota legal, cooperaciones con ONGs con proyectos activos. + Creación de las posibilidades extras para la incorporación de trabajadores Programas específicos para grupos discapacitados, inherente en la estrategia de empresa		

2.3.2. C3 Promoción del Comportamiento Ecológico de las Personas Empleadas

El indicador C3 mide la preocupación de la empresa en la promoción de los comportamientos ecológicos dentro de la empresa. Tiene en cuenta varios aspectos desde formación a incentivos todos en el sentido de reducir la huella ecológica de la empresa y las personas trabajadoras. Esta es su tabla:

Criterios	Principiante	Avanzado	Experimentado	Ejemplar
Cocina/cantina de la empresa / alimentación durante la jornada laboral	Primeras medidas para la promoción de modelos sostenibles de alimentación (p. ej. opción vegetal)	Clara confesión de la empresa hacia unas prácticas alimenticias más sostenibles (reducción relevante de productos animales)	Alimentación mayoritariamente vegetariana- vegana + alimentos de origen local, estacional y biológico/orgánico	Alimentación en gran parte vegetariana-vegana + alimentos mayoritariamente de origen local, estacional y biológico/orgánico
Movilidad al puesto de trabajo: sistema de incentivación / comportamiento real	Primeras medidas para una política de movilidad sostenible (p. ej. sistemas de incentivación financieros para la utilización de transporte público, política de coche de empresa fijada por escrito: <130 g CO ₂ / Km.)	Política de movilidad sostenible consecuente (p. ej. cuando no hay ningún medio de transporte urbano público disponible, Car Sharing activo, aparcamientos de empleados sólo para Car Sharing)	La mayoría de los empleados utilizan el transporte urbano público / autobús / tren / bici / Car Sharing	Casi todos los trabajadores utilizan el transporte urbano público / bici / Car Sharing
Cultura de la organización, concienciación y procesos de empresa internos	Tematización puntual de aspectos ecológicos (p. ej.: boletín de noticias, etc.) La gerencia vive un comportamiento ecológico (p. ej. ningún coche de empresa desmesurado de prestigio)	Integración puntual de aspectos ecológicos en los programas de formación continua Lo trabajadores serán integrados en asuntos ecológicos (tematización regular, eventos informativos, etc.)	Integración regular de aspectos ecológicos en los programas de formación continua Lo trabajadores serán integrados regularmente en asuntos ecológicos	Programas de concienciación para todos los trabajadores (p.ej. toma de datos regular para tematizar el comportamiento ecológico, talleres sobre la huella de carbono) Medidas innovadoras como p. ej. "servicios verdes sociales"
Huella ecológica (de Carbono) de los trabajadores (EUP / PYME) ³	< 5 ha / empleado/a	< 4 ha / empleado/a	< 3 ha / empleado/a	< 1,8 ha / empleado/a

2.4. Machine Learnig

"El aprendizaje automático o aprendizaje de máquinas (del inglés, "Machine Learning") es el subcampo de las ciencias de la computación y una rama de la inteligencia artificial cuyo objetivo es desarrollar técnicas que permitan a las computadoras aprender. De forma más concreta, se trata de crear programas capaces de generalizar comportamientos a partir de una información suministrada en forma de ejemplos."³

Esta ciencia busca descubrir la manera de construir sistemas que vayan creciendo y aprendiendo con la experiencia. Por resumir o sintetizar el proceso, el aprendizaje automático trata de, dada una tarea en particular y un tipo de algoritmo, trata de mejorar el rendimiento de esta tarea en función de la experiencia que recibe en forma de ejemplos.

Se considera el aprendizaje automático como una mezcla entre ingeniería de computación y estadística. Mientras que con ingeniería de computación se tiene la capacidad de obtener modelos para resolver problemas y encontrar cuales son tratables o intratables, mediante estadística se puede inferir mediante conjuntos de datos distintos modelos de los problemas. De tal forma que mediante la unión de ambas se consigue centrar la atención en como los sistemas serían capaces de programarse ellos mismos.

Como se puede apreciar es un cometido amplio y multidisciplinar que abarca numerosas aplicaciones. Desde aplicaciones en el campo de la medicina para tratar con numerosos datos hasta la investigación en astronomía en el análisis de millones de instantáneas de astros estelares. Se aplica con éxito en muchos ámbitos de la investigación y el desarrollo. Una de sus aplicaciones es el reconocimiento de voz, numerosos sistemas actuales emplean el aprendizaje automático para obtener cada vez mejores resultados adaptados al usuario y fieles a la realidad. Otra de sus aplicaciones es la de procesamiento de imágenes. En el campo de la astronomía se emplea machine learning para tratar con millones de fotografías tomadas del espacio y tratar de obtener clasificaciones. Si se empleasen medios convencionales esta tarea, aparte de muy laboriosa sería casi imposible debido al tiempo que se tardaría en analizar tanta información.

Dos características de problemas donde se puede aplicar este tipo de tecnología responden a la complejidad y la adaptación. En sistemas donde el desarrollo del algoritmo se vuelve demasiado complejo, por ejemplo, por el número de variables, se recurre a este tipo de tecnología puesto que reduce el tiempo para encontrar un algoritmo útil para cantidades enormes de datos de entrenamiento y que a priori no se ven sus similitudes. En segundo lugar, se emplea para momento en los que el software necesita adaptarse al lugar donde va a operar. Se busca que el sistema, una vez desarrollado, mejore en el ambiente donde se va a ejecutar para proporcionar una

³ Wikipedia. Aprendizaje Automático [En línea]
https://es.wikipedia.org/wiki/Aprendizaje_autom%C3%A1tico

mejor experiencia a la persona. Este es el caso típico de los sistemas de reconocimiento de voz, que necesitan adaptarse al usuario e ir mejorando para dar mejores resultados.

Esta tecnología está muy presente en el mundo actual y la podemos encontrar en numerosas aplicaciones desde las muy simples hasta las muy complejas. Pero a pesar de que este al alza sigue dando preguntas a las que es difícil responder, tanto en la parte técnica como en la ética. Una de ellas es si se pueden construir sistemas de aprendizaje continuo. La realidad es que los algoritmos se ejecutan sobre un conjunto de datos dado y obtiene unos resultados aplicables y centrado en el caso que le refiere, pero en la naturaleza el proceso nunca termina y se sigue aprendiendo indefinidamente y en variedad de ámbitos. Otra de las incógnitas es si el aprendizaje automático nos puede ayudar a entender cómo se lleva a cabo el proceso de aprendizaje humano. Por ejemplo, con las redes neuronales.

Siguiendo con las preguntas que nos arroja esta tecnología, ¿podemos diseñar un lenguaje de programación que autogenera? Es decir, ¿se puede crear un programa inacabado que termine de programarse el solo? Muchas son las dudas que quedan por resolver, pero sin duda es un campo que no deja de crecer y que va desarrollando nuevas aplicaciones. Recientemente se ha trabajado en un programa que va escribiendo un conocido libro a partir de sus entregas anteriores. Usando como datos de entrenamiento los libros anteriores y mediante procesadores de texto avanzado está escribiendo la última entrega. Obviamente no está dando los mejores resultados puesto que no es un humano el que lo escribe, pero es cierto que abre un mundo de posibilidades.

3.Propuesta de Estudio

3.1. Planteamiento del Problema

Como se puede observar a lo largo de la exposición hay varios problemas que saltan a la vista y que permiten buscar opciones de mejora. Estos son los problemas que he visto durante el estudio del contexto:

- **El proceso de creación de desarrollo del balance se torna lento y pesado.** Requiere un análisis pormenorizado de todos los aspectos de la empresa y además hace falta tener en cuenta numerosos factores para cada valoración. Esto hace que dure en ocasiones 100 horas para conseguir realizar una primera valoración. A partir del primer balance el resto duran menos. Es cierto que hay empresa que tienen esa capacidad, pero hay otras que no siempre pueden invertir tanto tiempo en realizar el estudio. Este problema temporal también se aplica a las personas que auditan. Tiene que invertir buena parte de su tiempo en estudiar las empresas. Además, aunque la EBC sea estatal no tiene filiales en todas las regiones de España. Por lo tanto, hay zonas donde no siempre va a haber personas capaces de llevar el proceso.
- **El proceso de valoración tiende a ser subjetivo y sufre variaciones en las distintas empresas.** Como se puede observar a través de las tablas internas los criterios y los niveles asociados no están definidos de forma extremadamente precisa. Es más ciertos niveles son demasiado subjetivos y se prestan a valoraciones personales. De esta forma se producen valoraciones que difieren mucho entre empresas, aunque a priori tengan características parecidas y tuvieran que tender la misma puntuación. Por ejemplo, tomando de referente el indicador A1, para valorar el nivel avanzado de aspectos regionales no se matiza cuanto es “algunos productos”. ¿Valdría con uno? ¿Hace falta que sea el 20% de las compras? No termina de quedar claro y se presta a la interpretación de la persona que evalúa.
- **Para realizar el balance se requiere de una formación específica.** A la hora de realizar el balance se requiere de cierta formación que te permita evaluar tu contexto empresarial y evaluar en consecuencia. Es cierto que esta formación se necesita para elaborar un informe detallado, pero a lo mejor no tendría que hacer falta para realizar una valoración inicial que sirva de borrador.

En definitiva, los problemas que achaco a este tipo de valoración son **TIEMPO, SUBJETIVIDAD Y FORMACIÓN.**

3.2. Propuesta de Solución

Vistos los problemas en la evaluación se propone el estudio de una solución que intente paliarlos. La idea es trabajar en un modelo de predicción de las evaluaciones. Una serie de ecuaciones o de árboles de decisión que sirvan para puntuar a las empresas, al menos en un momento inicial, sin la necesidad de un proceso tan extenso.

Puesto que se dispone de documentos de evaluación ya elaborados, la idea es trabajar sobre estos. El sistema de evaluación es subjetivo y, pero se dispone las auditorias previas. Por lo tanto, la propuesta es analizar cada documento e intentar estudiar todas las medidas que contiene para un determinado atributo y ver que puntuación han obtenido. Este análisis ha de ser manual, puesto que hay que inferir datos que no siempre esta escritos. Se podría intentar con algún procesador de lenguaje, pero se perdería información.

Una vez analizados los documentos, se dispondrá de un conjunto de datos que servirán como entrenamiento. Al tener datos de entrenamiento se prodrá trabajar con ellos mediante procesos de minería de datos y aprendizaje automático. El objetivo de este estudio es encontrar el modelo a partir de estos datos de entrenamiento que nos sirva para futuros casos. Este modelo será una solución aproximada de la valoración final. A través de distintos algoritmos de Machine Learning se intentará encontrar el que mejor se adapte a la realidad.

¿Cómo afecta esta solución al problema? Es sencillo. Si se tienen ecuaciones que sirvan para obtener valoraciones a priori, muchas empresas podrán dar el salto a este tipo de balances sin mucho esfuerzo. El tiempo que requiere estudiar la empresa y detallar cada una de medidas lo reducimos exponencialmente al tener una ecuación con unos atributos generales que sirvan para un primer acercamiento. Esta evaluación no va a ser la final, pero adelanta trabajo y permite empezar. En estudios posteriores cuando se disponga de más balances se puede abordar el problema de automatizar el proceso al completo. La subjetividad en la aplicación se reduce puesto que a partir de los ejemplos obtenemos un modelo en el que solo hay que rellenar campos para obtener la puntuación. Y dada la simpleza de este tipo de herramienta no hace falta ninguna formación previa para obtener el balance de tu empresa y comenzar a participar en la red de empresas del bien común.

En definitiva, si se consigue obtener conclusiones a partir de todos los datos recogidos de los balances se puede haber conseguido una primera aproximación para la automatización del proceso de evaluación de empresas. Esto supondría un avance cualitativo para esta iniciativa puesto que ahorraría costes, tiempo y facilitaría el proceso en general.

4. Investigación

A lo largo de este punto se va a recorrer todo el proceso de investigación que se ha seguido para la consecución del proyecto.

4.2. Modelación de datos

En este apartado se realizará una explicación del proceso seguido para recoger los datos, el modelado de los datos, las dificultades encontradas y todo lo referente a la recogida de datos.

Lo primero que se hizo en el proceso de recogido de datos es reflexionar mientras se estudiaban los documentos que la economía del bien común había facilitado. Había mucha información que procesar y por lo tanto había que intentar sistematizar todo para que fuera manejable.

Se tomó la decisión de modelar solo algunos indicadores de la matriz. Esto fue así puesto que para haber estudiado los 17 atributos hubiera requerido un esfuerzo demasiado grande al tener que analizar todo manualmente. Por lo tanto, se decidió centrarse en unos pocos y sacar conclusiones a partir de ellos. Los tres Indicadores elegidos fueron el A1, C1 y C3. El A1 se escogió por ser uno de los más amplios, puesto que le afectaban a priori numerosos factores. El C3 se escogió por tener muchas variables. Y el C3 por ser simple.

El proceso de extracción de datos lento y minucioso. Se leyeron los documentos que contenían las auditorías de las empresas que se habían realizado hasta la fecha. Como solo se habían seleccionado unos pocos indicadores solo se centró la lectura en esos. Mediante el estudio en profundidad y la comparación se obtuvieron iniciativas, acciones o procedimientos que en principio parecían afectar a las valoraciones finales. Las iniciativas encontradas en unas auditorías se cruzaban con otras para elaborar un conjunto de datos completo. A la vez se iba estudiando su relación o no con los niveles que a priori aporta la documentación que enseña como valorar compañías.

Con todo esto se construyó un listado y se dotó a los atributos con opciones definidas. De tal forma que se preparó un conjunto de atributos hipótesis que determinarían la puntuación del indicador. Este modelo es el que se ha empleado para investigar y partir del cual se han obtenido las ecuaciones de predicciones.

Todos los atributos obtenidos son nominales excepto la salida que es un numero entero, un porcentaje.

Uno de los puntos críticos a la hora de describir los atributos ha sido evitar la subjetividad. La mayoría se ha conseguido, pero en otros no habido más opción que introducir valores que están sometidos a la visión del que valora.

A continuación, se expondrán los modelos.

4.1.1. Modelo A1 Gestión ética de los suministros

Nombre	Valores	Pequeña descripción
Tamaño	Autónomo, reducido, medio, grande	Este atributo responde al conjunto de entrenamiento. Se dispone de datos de empresas de 1 empleado, de 2 a 15 y de 250.
NATURALEZA, PROPIEDADES Y CARACTERÍSTICAS DE LOS SUMINISTROS		
Móvil	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios móviles, incluye los terminales.
Fijo	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios de telefonía e internet.
Agua	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios de agua, incluye acciones para evitar gasto
Electricidad	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios de electricidad, incluye propia producción.
Banca	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios de banca y donde se tiene los depósitos.
Gestoría	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los servicios de gestión de la empresa, se aplica sobre todo a pequeñas empresas que tienen que externalizar.
Prov. Servicios	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del resto de proveedores con los que se contratan los servicios.
Prov. Material	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del resto de proveedores a los que se compran materiales para la producción.
Seguros	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del proveedor con el que se contratan los seguros necesarios para el negocio.
Transporte	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local, Público	Naturaleza del transporte utilizado en el desempeño de la actividad comercial.

Material Ofic	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza del material de oficina que se utiliza en la compañía. Incluye papeles, tinta, muebles, equipos...
Material	Desconocido, Convencional, Preocupación, Ético, Ecológico, Local	Naturaleza de la materia prima empleado en la producción.
Ahorro Ecológico	Si, No	Si se observan acciones enfocadas ecológicos en el proceso de compra a los proveedores
Eficiencia Energética	Si, No	Si se observan acciones enfocadas a reducir el consumo energético y aumentar la eficiencia de este.
Reciclaje	Si, No	Si se observan acciones de reciclado en la adquisición de suministros.
MEDIDAS ACTIVAS A LA HORA DE COMPRAR LOS SUMINISTROS		
Iniciativa	Muy baja, baja, media, alta, Muy alta	Nivel de iniciativa y de implicación que muestra la empresa a la hora de adquirir suministro
Laboral	Si, No	Si se aprecia preocupación por las condiciones laborales de las empresas proveedoras
Rechazo ético	Si, No	Si se aprecia rechazo a proveedores que no cumplan sus expectativas éticas
Abuso Prov	Si, No	Si se aprecia alguno abuso en la cadena de creación de valor
Formación	Si, No	Si hay acciones enfocadas a la formación para la mejora de la empresa en términos de la EBC
Estudio Impacto	Si, No	Si se ha realizado estudios del impacto de los proveedores en el entorno inmediato
Encuestas	Si, No	Si se realizan encuestas proveedores
MARCO ESTRUCTURAL PARA UN PRECIO JUSTO.		
Compras Locales	Ninguna, parcial, mayoría, todas	Nivel de compras a proveedores en las proximidades
Proveedores Fijos	Si, No	Si mantiene los proveedores o si cambian constantemente. Esto se valora en función de los años de actividad y la duración de sus contratos de suministros

Familiaridad Proveedores	Si, No	Si el trato con los proveedores es familiar, de cercanía y transparente
Concursos	Si, No	Si se realizan concursos o se recurre a subastas para contratar
Precios Ajustados	Si, No	Si los precios se ajustan en función de los proveedores y no en función del mercado, lo que implica que no fuerza al resto de implicados en la cadena de valor.

Hay pequeñas aclaraciones que hace falta aportar.

En primer lugar, aclarar cómo se han valorado los sí y los no en el proceso de recogida de datos. Es simple, si se aprecia alguna iniciativa en ese sentido o se hace alguna apreciación se ha marcado con un sí. En caso contrario, si no hay referencias o no se aprecian acciones, se marca como un no.

Otro de los puntos críticos que se comentó en el punto anterior es el de la subjetividad de algunos puntos. Los atributos Iniciativa y compras locales no se han podido establecer de otra forma puesto que no se conocen las cantidades exactas.

Los desconocidos en los primeros atributos vienen porque no siempre aparece la información reflejada en los balances. De esta forma se hace constar que el atributo no se conoce, y aunque a priori debería marcarse como convencional. Pero para entrenar lo marcaremos como desconocido para no dar por supuesto cosas que se desconocen.

En principio según se ha observado hay datos de los que vamos a prescindir en un futuro. Pero como ya se ha dicho esto es un prototipo inicial y han de indicarse.

Para cerrar este punto habría que señalar uno de los problemas en la recogida de datos de este atributo. Como ya se indicó al principio del documento existe una técnica que aconseja la EBC para recoger los datos de los proveedores. Lo que indican es que se realice una tabla con los pesos de cada uno de los proveedores y se indique el impacto ecológico y ético de cada uno pondera con los demás. Una tabla como esta:

Proveïdor	Proximitat	Impacte alt risc	Precio justo	Pes facturació	Proximitat pond	Impacte ecològic ponderat	Preu just ponderat
Proveïdor 1	28	0	0	11,08%	3,05	-	-
Proveïdor 2	55	30	10	8,76%	4,82	2,63	0,88
Proveïdor 3	9	19	0	7,05%	0,65	1,32	-
Proveïdor 4	55	19	0	6,77%	3,72	1,27	-
Proveïdor 5	55	15	10	6,29%	3,46	0,94	0,63
Proveïdor 6	46	15	10	5,46%	2,5	0,82	0,55
Proveïdor 7	9	30	7	4,17%	0,38	1,25	0,29
Proveïdor 8	9	8	0	3,22%	0,3	0,24	-
Proveïdor 9	28	30	15	3,02%	0,83	0,91	0,45
Proveïdor 10	9	15	7	3,01%	0,28	0,45	0,21
Proveïdor 11	18	11	0	2,93%	0,54	0,33	-
Proveïdor 12	55	23	0	2,81%	1,55	0,63	-
Proveïdor 13	0	0	0	2,78%	-	-	-
Proveïdor 14	37	30	10	2,52%	0,92	0,76	0,25
Proveïdor 15	9	0	0	2,21%	0,2	-	-

Proveïdor 55	28	11	0	0,23%	0,06	0,03	-
Proveïdor 56	55	26	0	0,20%	0,11	0,05	-
Proveïdor 57	0	0	0	0,19%	-	-	-
Proveïdor 58	9	30	0	0,19%	0,02	0,06	-
Proveïdor 59	28	26	0	0,17%	0,05	0,04	-
Proveïdor 60	0	26	0	0,16%	-	0,04	-
Proveïdor 61	0	30	0	0,14%	-	0,04	-
Proveïdor 62	55	11	0	0,13%	0,07	0,01	-
Proveïdor 63	9	30	0	0,11%	0,01	0,03	-
Proveïdor 64	0	30	0	0,11%	-	0,03	-
Proveïdor 65	55	15	0	0,11%	0,06	0,02	-
Proveïdor 66	0	30	0	0,07%	-	0,02	-
Proveïdor 67	55	19	0	0,06%	0,04	0,01	-
Proveïdor 68	0	30	0	0,03%	-	0,01	-
Total				100%	31,51	16,05	3,63
				% indicador :	57,30%	53,50%	24,20%

Como se puede apreciar este tipo de estudio puede ayudar y mucho a la modelización. El problema es que muy pocas empresas lo han realizado. En un principio se pensó en realizar los estudios de los que faltaban, pero no se disponía de la suficiente información y de descarto la idea. Se decidió ignorar estas tablas y obtener la información de otras partes de los balances.

4.1.2. Modelo C1 Calidad del Puesto de Trabajo e Igualdad

Nombre	Valores	Pequeña descripción
Tamaño	Autónomo, reducido, medio, grande	Este atributo responde al conjunto de entrenamiento. Se dispone de datos de empresas de 1 empleado, de 2 a 15 y de 250.
Horario de trabajo		
Horario	Trabajador, Proyectos, Establecido, Democrático	Forma en la que se establece el horario de trabajo de las personas empleadas en la empresa.
Flexibilidad	Autogestión, personalizada, familiar, ninguna, NA	Dentro del horario de trabajo cuando de flexible es una vez establecido.
Días libres	Establecido, libre, adaptado, NA	Como se reparten los días libres en la empresa.
Vacaciones	Establecido, libre, adaptado, NA	Como se reparten las vacaciones en la empresa.
Configuración del puesto de trabajo		
Lugar	Oficina, hogar, mixto, trabajador	Donde se establece el lugar de trabajo.
Fenshui	Si no	Si se tienen en consideración medidas de fenshui en la organización espacial de la empresa.
Zonas verdes	Si no	Existencia o no de zonas verdes en el lugar de trabajo.
Accesibilidad	Si no	Accesibilidad a personas con movilidad reducida en el puesto de trabajo.
Solicitudes	Si no	Si existe un método para realizar solicitudes o peticiones. de cambio o mejor en el lugar de trabajo.
Ergonomía	Si no	Si se dispone de mobiliario ergonómico.
Espacios amplios	Si no	Si se dispone de espacios amplios donde descansar y estar en el lugar de trabajo.
Ventilación	Si no	Si dispone de un sistema de ventilación particular.
Salud Física y Seguridad		
Previsión riesgos	si no	Si la empresa contempla algún sistema de previsión de riesgos.
Revisiones medicas	si no	Si la empresa tiene previstas revisiones médicas periódicas.

Comida	Cooperativa, trabajador, empresa	Quien provee la comida dentro de la empresa.
Formación salud	si no	Si las trabajadoras reciben algún tipo de formación en temas de salud.
Relajación	si no	Si se imparte formación en relajación o se realizan actividades relacionadas.
Estudio Seguridad	si no	Si se ha realizado algún estudio de seguridad.
Actividades y talleres	si no	Si se realizan actividades y talleres.
Seguro	si no	Si se dispone de algún tipo de seguro especial para las empleadas.
Salud Psíquica		
Formación	Empresa, libre, Facilitado	A quien correo cargo la formación en temas de salud mental.
Clima Bueno	si no	Si la empresa goza de un buen clima de trabajo.
Actividades desarrollo personal	si no	Si la empresa organiza actividades dedicadas al desarrollo personal.
Encuestas valorativas	si no	Si disponen de encuestas periódicas para valorar el estado de las personas que allí trabajan.
Organización propia, satisfacción en el puesto de trabajo, sentido		
Decisiones	Jerárquico, horizontal, consensuadas	Como se toman las decisiones en la empresa.
Programas de conciliación	si no	Si existen programas específicos dedicadas a la conciliación familiar.
Herramientas Organizativas	si no	Si existen herramientas dedicadas a la organización de la empresa.
Medidas Comunicación	si no	Si existen medidas o planes dedicados a la comunicación interna en la empresa.
Encuestas satisfacción	si no	Si existen encuestas de satisfacción.
Igualdad y trato igualitario para hombre y mujer		
Remuneración	Igualitaria, positiva, discriminatoria, NA	Como se establecen las remuneraciones en función al género.

proporción	0, <25,25-50,50,50-75,>75, NA	Cuál es la proporción en el género. En función de las mujeres que haya.
Lenguaje Inclusivo	si no NA	Se utiliza lenguaje inclusivo en los documentos de la empresa.
Talleres Formativos	si no NA	Si existen talleres formativos para temas de género.
Plan de inclusión	si no NA	Si hay algún de corrección de género.
Medidas Preventivas	si no NA	Si existen medidas para prevenir actitudes machistas.
Desfavorecidos		
Alumnos	si no NA	Si tiene un programa de prácticas para estudiantes.
Desempleados	si no NA	Si tiene un programa de contratación de parados de larga duración.
Inmigrantes	si no NA	Si tiene un programa de contratación de inmigrantes.
movilidad reducida	si no NA	Si tiene un programa de contratación para personas con movilidad reducida.
diversidad funcional	si no NA	Si tiene un programa de contratación para personas con diversidad funcional.
exclusión	si no NA	Si tiene un programa de contratación para personas en situación de exclusión social.

Como se hizo en el anterior punto hacen falta algunas aclaraciones. Se puede apreciar que el número de variables es sensiblemente mayor, y es que en este caso el indicador es mucho más extenso y abarca más campos.

Es punto ha dado numerosos problemas. Al ser tan extenso es complicado obtener los puntos que se supone influirán en la puntuación final. No en todas las empresas se seguían el mismo patrón y las mismas condiciones. De hecho, se ha decido optar por introducir el valor NA (No Aplica) a las variables que solo estuvieran presentes en algunas empresas. El principal problema de este indicador, como ya se verá en la experimentación es que las valoraciones no se ajustan a la realidad. Hay muchas empresas sobrevaloradas y que no siguen los criterios como indica el manual de calificación. Eso hace que empresas con bajas cualidad se otorguen grandes puntuaciones y, en cambio, otras con grandes cualidades obtengan puntuaciones mediocres.

4.1.3. Modelo C3 Promoción del comportamiento ecológico de las personas empleadas

Nombre	Valores	Pequeña descripción
Tamaño	Autónomo, reducido, medio, grande	Este atributo responde al conjunto de entrenamiento. Se dispone de datos de empresas de 1 empleado, de 2 a 15 y de 250.
Cocina/Cantina de la empresa/Alimentación		
Lugar	Casa, restaurante, oficina	Indica donde comen las personas que trabajan en la empresa.
vegetariano	Ninguna, preocupación, mayoría, todo	Cuántas alternativas vegetarianas se dan en la comida de la empresa, ya sea servicio propio o por del empleado.
cocina	Si no	Si la empresa dispone de cocina para las personas que trabajan allí o no
comida	Ecológica, local y convencional	Procedencia de la comida que se consume en la empresa, ya sea servicio propio o por parte del empleado.
Formación	Si no	Si se imparte formación sobre temas alimenticios y nutricionales.
office	Si no	Si se dispone de una sala para comer.
Proyecto Colab.	Si no	Si hay algún proyecto colaborativo con otra entidad para la comida de la empresa.
Restaurante oficina	Si no	Si existe en la empresa servicio de restauración.
Movilidad al puesto de trabajo: sistema de incentiviación / comportamiento real		
Llegada	ninguna, caminando, bicicleta, público, compartido, empresa, privado	Como acceden la mayoría de las personas de la empresa a su puesto de trabaja.
Desplazamientos	ninguna, caminando, bicicleta, público, compartido, empresa, privado	Como se realizan los desplazamientos dentro del horario de trabajo.
Ubicación	Centro, fuera	Donde se ubica la empresa en relación al caso urbano más cercano.
Contratación Cercana	Si no	Si se prioriza la contratación de personas próximas al lugar de trabajo.
vestuario	Si no	Si se dispone de vestuario dentro de la empresa.

Reuniones Virtuales	Si no	Si a menudo se realizan reuniones virtuales para reducir los desplazamientos.
Facilidad bici	Si no	Si existen facilidades para acceder en bicicleta.
Facilidad Publico	Si no	Si existen facilidades para acceder en transporte público.
Facilidad Aparcamiento	Si no	Si existen facilidades para acceder en coche al puesto de trabajo.
Apoyo financiero	Si no	Si hay apoyo financiero para utilización de transportes alternativos al privado.
Concienciación	Si no	Si existen mecanismos de concienciación para los empleados de la empresa
Ejemplaridad	Si no	Si se da ejemplo a otras empresas de su alrededor
Incentivos	Si no	Si se incentiva a las personas que trabajan allí para favorecer comportamientos ecológicos.
Material	Si no	Si disponen de materia para concienciar.
Practicas eco	Si no	Si han implantado prácticas ecológicas que repercutan en su huella ecológica.
Formación	Si no	Si se imparta formación sobre la reducción de la huella ecológico.
Directivos	ninguna, caminando, bicicleta, publico, compartido, empresa, privado	Como se desplazan las personas en la cima de la empresa.
Análisis	Si no	Si se ha realizado un análisis sobre la contaminación de la empresa.
Huella Ecológica		
Huella	'<5','<4','<3','<1.8', ninguna	

4.2. Algoritmos investigación

Llegados a este punto, se van a mostrar los algoritmos empleados en la experimentación. Cada algoritmo irá acompañado de una pequeña historia y en los casos donde sea posible su pseudocódigo.

4.2.1. Funciones

4.2.1.1. Perceptrón Multicapa

Este algoritmo es una generalización del Perceptrón simple y se creó a raíz de los límites en la resolución de problemas no lineales. Se le considera el aproximador universal y es de los algoritmos más empleados por esta razón.

Su arquitectura se basa en imitar un conjunto de neuronas agrupadas en capas. Existen tres tipos de capas en las que se pueden organizar estas: entrada, oculta y salida. En la primera, la de entrada, actuaría como receptores para recibir los estímulos externos. Todas estas neuronas se conectan, habitualmente, a la siguiente capa de manera cruzada con multitud de conexiones. Una vez la información ha pasado por las distintas capas oculta se llega a las de neuronas de salida en las que se muestra el resultado final.

Cada una de estas conexiones tiene un valor específico y es la que da como resultado las variables de salida en función de la entrada determinada. Cada neurona de la capa oculta procesa la información recibida y la pasa de capa en capa hasta llegar a la salida. El algoritmo generalizado de aprendizaje sigue los siguientes pasos:

1. Se inicializan los pesos de la red. Se hace de manera aleatoria y con pesos cercanos a 0.
2. Se toma un ejemplo del conjunto de entrenamiento y se propaga. Las neuronas de entrada se activan mediante:

$$a_i^1 = x_i \text{ para } i = 1, 2, \dots, n_1$$

Donde $X = (x_1, x_2, \dots, x_{n_1})$ representa el vector o patrón de entrada a la red.

Las neuronas de la capa oculta lo hacen aplicando la función de activación f y procesando todos los datos que han recibido de la capa de entrada ponderados por los pesos de las conexiones correspondientes. Su ecuación sería la siguiente:

$$a_i^c = f \left(\sum_{j=1}^{n_{c-1}} w_{ji}^{c-1} a_j^{c-1} + u_i^c \right) \text{ para } i = 1, 2, \dots, n_c \text{ y } c = 2, 3, \dots, C-1$$

donde a_j^{c-1} son las activaciones de la capa $c-1$

A continuación, las salidas de las neuronas de la capa de salida vienen determinadas por la capa anterior, la última de las capas ocultas. En general:

$$y_i = a_i^C = f \left(\sum_{j=1}^{n_{C-1}} w_{ji}^{C-1} a_j^{C-1} + u_i^C \right) \text{ para } i = 1, 2, \dots, n_c$$

donde Y es el vector de salida de la red

3. Se evalúa el error de la salida.
4. Se aplica la regla delta para modificar los pesos y umbrales de la red:

- 4.1. Se calculan los valores de δ para todas las neuronas de la capa de salida utilizando, Sea w_{ji}^{C-1} el peso de la conexión de la neurona j de la capa $C - 1$ a la neurona i de l

$$\delta_i^C(n) = -(s_i(n) - y_i(n))f' \left(\sum_{j=1}^{n_{C-1}} w_{ji}^{C-1} a_j^{C-1} + u_i^C \right)$$

- 4.2. Se calculan ahora para los de las siguientes capas, de atrás hacia las de entrada, mediante:

$$\delta_j^{c+1}(n) = f' \left(\sum_{k=1}^{n_c} w_{kj}^c a_k^c + u_j^c \right) \sum_{i=1}^{n_{c+1}} \delta_i^{c+2}(n) w_{ji}^c$$

donde $a_k^c(n)$ es la activación de la neurona k de la capa c para el patrón n

- 4.3. Se modifican los pesos y los umbrales mediante:

$$w_{ji}^{C-1}(n) = w_{ji}^{C-1}(n-1) + \alpha \delta_i^C(n) a_j^{C-1}$$

$$\text{para } j = 1, 2, \dots, n_{C-1} \text{ } i = 1, 2, \dots, n_C$$

$$u_i^C(n) = u_i^C(n-1) + \alpha \delta_i^C(n) \text{ para } i = 1, 2, \dots, n_C$$

para la capa de salida, y con:

$$w_{kj}^c(n) = w_{kj}^c(n-1) + \alpha \delta_j^{c+1}(n) a_k^c(n)$$

$$\text{para } k = 1, 2, \dots, n_c, \text{ } j = 1, 2, \dots, n_{c+1} \text{ y } c = 1, 2, \dots, C-2$$

$$u_j^{c+1}(n) = u_j^{c+1}(n-1) + \alpha \delta_j^{c+1}(n)$$

$$\text{para } j = 1, 2, \dots, n_{c+1} \text{ y } c = 1, 2, \dots, C-2$$

para el resto de capas de la red

5. Se repite 2,3,4 para todos los ejemplos de entrenamientos.
6. Completado un ciclo de entrenamiento, se obtiene el error general.
7. Se repite del 1-6 hasta que se consigue minimizar el error general del entrenamiento.

Referencias⁴

⁴ Perceptron Multicapa [En línea] <http://bibing.us.es/proyectos/abreproy/12166/fichero/Volumen+1+-+Memoria+descriptiva+del+proyecto%252F3+-+Perceptron+multicapa.pdf>

Wikipedia. Perceptron [En línea] https://es.wikipedia.org/wiki/Perceptr%C3%B3n_multicapa

4.2.1.2. Regresión lineal

Este algoritmo trata de aproximar una función $f(x)$ de naturaleza lineal con una función lineal determinada. Su objetivo es minimizar el error entre la función estimada y el valor original.

La solución a este problema pasa por ir variando los pesos de la supuesta ecuación hasta conseguir una aproximación con el mismo error. De tal forma siendo la ecuación:

$$\hat{f}(\vec{x}) = w_0 + w_1 a_1(\vec{x}) + w_2 a_2(\vec{x}) + \dots + w_n a_n(\vec{x})$$

se trata de ir variando los pesos w hasta encontrar la ecuación óptima.

El método de aprendizaje se describe de la siguiente forma. Siendo $\{x, t\}$, una dupla formada por cada vector de entrenamiento (x) y las salidas reales (t), y n la ratio de entrenamiento:

1. Inicializar w a un valor cualquiera, a ser posible pequeño.
2. Hasta que el error sea minimizado
 - a. Inicializar cada Δw_i a cero.
 - b. Para cada $\{x, t\}$
 - i. Calcular el valor $o = f(x)$
 - ii. Para cada peso: $\Delta w_i \leftarrow \Delta w_i + \eta(t - o)x_i$
 - c. Para cada peso w : $w_i \leftarrow w_i + \Delta w_i$

Referencias ⁵

⁵UC3m. Aprendizaje automático <http://ocw.uc3m.es/ingenieria-informatica/aprendizaje-automatiko/material-de-clase-1/aa-ocw-regresion.pdf>

4.2.2. Árboles de decisión

4.2.2.2. C4.5

Este algoritmo se emplea para conseguir obtener árboles de decisión. Fue desarrollado por Ross Quinlan. Es una mejora del anterior sistema ID3 que desarrolló creó el mismo Ross.

Tiene el mismo principio que ID3 utiliza la entropía para clasificar los ejemplos de entrenamiento y obtener un árbol generalizado. En cada uno de los nodos se elige el atributo que mejor divida el conjunto o los subconjuntos si no es el primer paso. Esta decisión se toma en función de la diferencia de entropía entre atributos.

La entropía se explica de manera simple. Si se observa un atributo y los valores que tiene se puede dividir el conjunto en función de estos. Tendríamos los ejemplos que pertenecen a unos valores y los que no. La entropía mide esa proporción. Cuando un atributo divide los datos en conjuntos similares, la entropía sería máxima. Es decir si para un atributo con dos valores se obtiene 0.5 de cada uno, la entropía sería la máxima.

Por lo tanto sabemos que esa la entropía este algoritmo consiste en :
consiste en :

1. Comprueba los datos iniciales.
2. Para cada atributo
 - a. Encontrar la entropía
3. Escoger el mejor atributo, el que tenga la mejor entropía con respecto a los demás.
4. Crear el nodo en función de ese atributo.
5. Repetir 1-4 para cada subconjunto de datos generado por el nodo.

J48 y RandomTree no son más que adaptaciones en código de este algoritmo.

Referencias ⁶

⁶ Wikipedia. C4.5 [En línea] <https://es.wikipedia.org/wiki/C4.5>

4.3. Entorno de Desarrollo

4.3.1. Hardware

Para la realización de esta práctica se han utilizado dos ordenadores personales. Uno de sobremesa y otro portátil. El primero dispone de las siguientes características:

- Procesador: AMD FX™-8320 Ocho núcleos 3.5 GHz
- Memoria RAM: 8GB
- Disco Duro: Disco Duro 1 T B y Disco SSD 256 GB
- Tarjeta Gráfica: Nvidia Gforce GTX 1060 6 GB
- Sistema Operativo: Windows 10 Pro, 64 bits

Más tarde se adquirió un ordenador portátil para el último mes de trabajo. Sus características son:

- Fabricante: Asus
- Procesador: Intel® Core™ i5-6300HQ Quad-Core (6M Cache, 2.3GHz hasta 3.2GHz)
- Memoria RAM: 8 GB
- Disco Duro: 1 TB
- Tarjeta Gráfica: Nvidia Gforce GTX 960
- Sistema Operativo: Windows 10 Pro, 64 bits

4.3.2. Software

El software elegido para el desarrollo de la investigación ha sido weka. Es un entorno en el que ya se han realizado prácticas anteriores y del que se dispone de conocimiento previo.

Weka⁷ (Waikato Enviroment for Knowledge Analisis) es una herramienta desarrollada por la universidad de Waikato. Es una aplicación de software libre que implemente funciones de aprendizaje automático y minería de datos.

Esta herramienta permite cargar los datos en ficheros. arff y se adapta a las necesidades de investigación. Contiene algoritmos de clasificación, regresión y clustering.

En ambos ordenadores se ha trabajado con la versión Pro de Windows 10.

⁷Wikipedia. Weka [En línea] [https://es.wikipedia.org/wiki/Weka_\(aprendizaje_autom%C3%A1tico\)](https://es.wikipedia.org/wiki/Weka_(aprendizaje_autom%C3%A1tico))

4.4. Pruebas a realizadas

En este punto se detallan las pruebas que se han realizado para cada uno de los modelos que se acaban de exponer en el punto anterior. Se definirán los conjuntos de datos que se han utilizado en las experimentaciones, es decir, los cambios que se han ido realizando en los modelos iniciales. Luego se presentará el total de las pruebas que se han hecho para cada atributo.

El proceso de pruebas ha sido metódico. Se probaban los algoritmos con los mismos parámetros para los distintos conjuntos de datos. Para aquellos que arrojaban mejores resultados se realizaban pruebas más exhaustivas, en cambio para los otros bastaba con unas pocas. Muchas veces cuando se encontraba el mejor algoritmo se realizaban pruebas rápidas con distintas modificaciones para encontrar el mejor conjunto de datos. Esto permitía encontrar mejores conjuntos de datos.

Aquí se ponen las pruebas recogidas, aunque ha habido más. Muchas pequeñas pruebas que no arrojaban resultados notables o que empeoraban. Se trabajó con distintos preprocesados de datos a parte de los empleados en eliminar variables, pero se descartaron. Por ejemplo, normalización o conversión a binario las variables nominales

4.4.1 Conjuntos de Datos A1

- **DatosA1:** Este conjunto de datos representa el modelo que inicial que se describió. Todos los atributos son nominales y tiene los valores que se han detallado. La clase también es nominal, agrupa los posibles puntos finales en grupos de 5 en 5 de 0 a 100, puesto que representa un porcentaje.
- **DatosA1_datos_eco,ético,local:** Supone una modificación del modelo inicial. Esta vez se resumen las primeras variables en 3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local y ético. Estas variables son numéricas y tiene valores del 1 al 5.
- **DatosA1_datos_eco,ético,local,preoc:** Supone una modificación del modelo inicial. Esta vez se resumen las primeras variables en 3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local, preocupación y ético. Estas variables son numéricas y tiene valores del 1 al 5.
- **DatosA1_clase_reducida:** Este conjunto de datos representa el modelo que inicial con una modificación en la clase. Todos los atributos son nominales y tiene los valores que se han detallado. La clase también es nominal, pero agrupa los posibles puntos finales en grupos de 10 en 10 de 0 a 100.
- **DatosA1_clase_reducida_datos_eco,ético,local:** Supone una modificación del modelo reducido. Se resumen las primeras variables en

3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local y ético. Estas variables son numéricas y tiene valores del 1 al 5.

- **DatosA1_clase_reducida_datos_eco,ético,local,preoc:** Supone una modificación del modelo reducido. Se resumen las primeras variables en 3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local, preocupación y ético. Estas variables son numéricas y tiene valores del 1 al 5.
- **DatosA1_clase_numérica:** Este conjunto de datos representa el modelo que inicial con una modificación en la clase. Todos los atributos son nominales y tiene los valores que se han detallado. La clase esta vez es numérica en vez de nominal.
- **DatosA1_clase_numérica_sinDesc:** Es el mismo modelo con la clase numérica solo que se eliminan todos los desconocidos por convencionales.
- **DatosA1_clase_numérica_datos_eco,ético,local:** Supone una modificación del modelo numérico. Se resumen las primeras variables en 3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local y ético. Estas variables son numéricas y tiene valores del 1 al 5.
- **DatosA1_clase_numérica_datos_eco,ético,local,preoc:** Supone una modificación del modelo numérico. Se resumen las primeras variables en 3. Todas aquellas que tenían como valores desconocido, convencional, preocupación, ético, ecológico o local se reducen a ecológico, local, preocupación y ético. Estas variables son numéricas y tiene valores del 1 al 5.
- **DatosA1_clase_numérica-iniciativa-movil:** Supone una modificación de la clase numérica. Se eliminan los atributos iniciativa y móvil.
- **DatosA1_clase_numérica-iniciativa-movil-abuso:** Supone una modificación de la clase numérica. Se eliminan los atributos iniciativa, móvil y abuso proveedores.

4.4.2. Pruebas A1

Datos A1	Inicial	Datos_eco, ético, local	datos_eco, ético, local, preoc
	Algoritmo: J48		
	CF: 0,01		
	CF:0,1		
	CF: 0,25		
	Algoritmo: Multilayer Perceptron		
	LR: 0,01 T: 500		
	LR: 0,01 T: 5000		
	LR: 0,1 T: 500		
	LR: 0,1 T: 5000		
	LR: 0,25 T: 500		
	LR: 0,25 T: 5000		
	Algoritmo:RandomTree		
	Seed:1		
	Seed: 5		
	Seed: 10		

Clase Reducida	Inicial	Datos _eco, ético, local	Datos _eco, ético, local, preoc
	Algoritmo: J48		
	CF: 0,01		
	CF:0,1		
	CF: 0,25		
	Algoritmo: Multilayer Perceptron		
	LR: 0,01 T: 500		
	LR: 0,01 T: 5000		
	LR: 0,1 T: 500		
	LR: 0,1 T: 5000		
	LR: 0,25 T: 500		
	LR: 0,25 T: 5000		
	Algoritmo: RandomTree		
	Seed:1		
	Seed: 5		
	Seed: 10		

Numérica	Inicial	Datos eco, ético, local	Datos eco, ético, local, preoc	SinDesc	-iniciativa -movil	-iniciativa -movil -abuso
	Algoritmo:LinearRegresion					
	Greedy					
	M5					
	NoAttrib					
	Algoritmo: Multilayer Perceptron				Algoritmo: Multilayer Perceptron	
	LR: 0,01 T: 500				LR: 0,01 T: 500	
	LR: 0,01 T: 5000					
	LR: 0,1 T: 500				LR: 0,1 T: 500	
	LR: 0,1 T: 5000					

4.4.3. Conjuntos de Datos C1

- **DatosC1:** Este conjunto de datos representa el modelo que inicial que se describió. Todos los atributos son nominales y tiene los valores que se han detallado. La clase también es nominal, agrupa los posibles puntos finales en grupos de 5 en 5 de 0 a 100, puesto que representa un porcentaje.
- **DatosC1_clase_numérica:** Este conjunto de datos representa el modelo que inicial con una modificación en la clase. Todos los atributos son nominales y tiene los valores que se han detallado. La clase esta vez es numérica en vez de nominal.
- **DatosC1_clase_reducida:** Este conjunto de datos representa el modelo que inicial con una modificación en la clase. Todos los atributos son nominales y tiene los valores que se han detallado. La clase también es nominal, pero agrupa los posibles puntos finales en grupos de 10 en 10 de 0 a 100.

4.4.4. Pruebas C1

Datos C1	Inicial	Reducida	Numérica
	Algoritmo: J48		Algoritmo:LinearRegresion
	CF: 0,01		Greedy
	CF:0,1		M5
	CF: 0,25		NoAttrib
	Algoritmo: Multilayer Perceptron		
	LR: 0,01 T: 500		
	LR: 0,01 T: 5000		
	LR: 0,1 T: 500		
	LR: 0,1 T: 5000		
	LR: 0,25 T: 500		
	LR: 0,25 T: 5000		
	Algoritmo: Random Tree		
	Seed:1		
	Seed: 5		
	Seed: 10		

4.4.5. Conjuntos de Datos C3

- **DatosC3:** Este conjunto de datos representa el modelo que inicial que se describió. Todos los atributos son nominales y tiene los valores que se han detallado. La clase también es nominal, agrupa los posibles puntos finales en grupos de 5 en 5 de 0 a 100, puesto que representa un porcentaje.
- **DatosC3_huellaNumeric:** Modificación del modelo inicial. Se sustituye la huella ecológica como variable nominal por una con variable numérica
- **DatosC3_huellaNumeric-huella-conc**
- **DatosC3_huellaNumeric-huella-conc-office**
- **DatosC3_huellaNumeric-huella-conc-office-fbici**
- **DatosC3_huellaNumeric-huella-conc-office-fbici-formacion**
- **DatosC3_huellaNumeric-huella-conc-office-fbici-formacion-analisis**
- **DatosC3_huellaNumeric-huella-conc-office-fbici-formacion-restaurante**
- **DatosC3_huellaNumeric-huella-conc-office-fbici-formacion-restaurante-analisis**
- **DatosC3_Mejor:** fichero de datos con el mejor resultado posible habiendo eliminado los atributos que daban mayores errores. Se han eliminado: huella, concienciación, office, fbici, formación, restaurante y análisis.

4.4.6. Pruebas C3

Datos C3	Inicial	Huella Numerica	NoAttribute	Mejor	Mejor Binary
	Algoritmo:LinearRegresion			Algoritmo: Linear Regresion	
	No Atribute		-Huella	No Atribute	
	M5		-Conciencia	M5	
	Greedy		-Office	Greedy	
	Algoritmo: Multilayer Perceptron		-Cocina	Algoritmo: Multilayer Perceptron	
	LR: 0,3 T: 500		-Fbici	LR: 0,3 T: 500	
	LR: 0,1 T: 500		-Formación	LR: 0,1 T: 500	
	LR: 0,01 T: 500		-Análisis	LR: 0,01 T: 500	
	LR: 0,3 T: 500 S:0	LR: 0,1 T: 500 S:0		LR: 0,1 T: 500 S:0	
	LR: 0,3 T: 500 S:5	LR: 0,1 T: 500 S:5		LR: 0,1 T: 500 S:5	
	LR: 0,3 T: 500 S:10	LR: 0,1 T: 500 S:10		LR: 0,1 T: 500 S:10	
	LR: 0,3 T: 500 S:15	LR: 0,1 T: 500 S:15		LR: 0,1 T: 500 S:15	
				Algoritmo: Tree	
				M5PTree	
				RandomTree	

5. Resultados

A lo largo de este punto se irán mostrando los resultados obtenidos a lo largo de la investigación. Se expondrá todo en tablas para facilitar la lectura. Todas las salidas se dan en torno al error relativo para que se pueda comparar entre sistema. Para todos los conjuntos de datos que tiene salida nominal viene acompañados de la tasa de aciertos puesto que es un dato que también aporta mucha información y con estos tenemos esa posibilidad. Con los numéricos no se aporta puesto que no se recibe del programa. Solo se mostrarán las ecuaciones de para el mejor resultado en cada caso, para no sobrecargar el documento. Las tablas son homólogas a las del punto anterior para mejorar su comprensión.

5.1. A1

Datos A1	Inicial	Datos_eco, ético, local	datos_eco, ético, local, preoc
Algoritmo: J48			
CF: 0,01	Ac: 21.4286 % Er: 92.9398 %		
CF:0,1	Ac: 21.4286 % Er: 89.1204 %		
CF: 0,25	Ac: 21.4286 % Er: 89.1204 %		
Algoritmo: Multilayer Perceptron			
LR: 0,01 T: 500	Ac: 21.4286 % Er: 94.0655 %		
LR: 0,01 T: 5000	Ac: 21.4286 % Er: 93.1869 %		
LR: 0,1 T: 500	Ac: 21.4286 % Er: 93.3576 %		
LR: 0,1 T: 5000	Ac: 21.4286 % Er: 91.1347 %		
LR: 0,25 T: 500	Ac: 21.4286 % Er: 92.7168 %		
LR: 0,25 T: 5000	Ac: 21.4286 % Er: 90.842 %		
Algoritmo: RandomTree			
Seed:1	Ac: 7.1429 % Er: 100.915 %	Ac: 21.4286 % Er: 84.1312 %	Ac: 28.5714 % Er: 76.4829 %
Seed: 5	Ac: 14.2857 % Er: 90.6869 %	Ac: 7.1429 % Er: 99.4278 %	Ac: 35.7143 % Er: 66.2852 %
Seed: 10	Ac: 14.2857 % Er: 95.6037 %	Ac: 42.8571 % Er: 61.1111 %	Ac: 14.2857 % Er: 91.7795 %

Datos A1	Inicial	Datos_eco, ético, local	datos_eco, ético, local, preoc
Algoritmo: J48			
CF: 0,01	Ac: 7.1429 % Er: 106.891 %		
CF:0,1	Ac: 7.1429 % Er: 103.8194 %		
CF: 0,25	Ac: 7.1429 % Er: 103.8194 %		
Algoritmo: Multilayer Perceptron			
LR: 0,01 T: 500	Ac: 14.2857 % Er: 98.505 %		
LR: 0,01 T: 5000	Ac: 7.1429 % Er: 101.4525 %		
LR: 0,1 T: 500	Ac: 7.1429 % Er: 101.4385 %		
LR: 0,1 T: 5000	Ac: 7.1429 % Er: 101.3718 %		
LR: 0,25 T: 500	Ac: 7.1429 % Er: 101.4973 %		
LR: 0,25 T: 5000	Ac: 7.1429 % Er: 101.5562 %		
Algoritmo: RandomTree			
Seed:1	Ac: 21.4286 % Er: 88.1812 %	Ac: 42.8571 % Er: 57.1429 %	Ac: 14.2857 % Er: 89.2301 %
Seed: 5	Ac: 7.1429 % Er: 92.8571 %	Ac: 21.4286 % Er: 87.1906 %	Ac: 21.4286 % Er: 84.1312 %
Seed: 10	Ac: 0 % Er: 110.4492 %	Ac: 28.5714 % Er: 76.4829 %	Ac: 21.4286 % Er: 83.0386 %

Numérica	Inicial	Datos eco, ético, local	Datos eco, ético, local, preoc	SinDesc	-iniciativa -movil	-iniciativa -movil -abuso
Algoritmo: Linear Regresion						
Greedy	39.5194 %	86.2112 %	71.8572 %	77.1009 %	29.7933 %	29.2837 %
M5	63.7925 %	67.472 %	83.3253 %	62.0877 %	57.7587 %	56.1126 %
NoAttrib	65.1734 %	71.8006 %	73.0054 %	62.6653 %	57.7216 %	58.5699 %
Algoritmo: Multilayer Perceptron						
LR: 0,01 T: 500	73.4409 %	65.4081 %			70.4995 %	
LR: 0,01 T: 5000	73.4566 %	63.817 %				
LR: 0,1 T: 500	71.9916 %	63.2511 %			69.3402 %	
LR: 0,1 T: 5000	71.8599 %	63.2511 %				

En primer lugar, analizar las diferencias entre los resultados con salida nominal y los de salida numérica. En primera instancia pensaba tratar la salida de manera nominal puesto que no hacía falta una salida exacta, bastaba con dar una aproximada entre un rango de valores. Pero según fueron avanzando las pruebas se intuía que era difícil obtener un error relativo reducido. Todo lo que se obtenían era aciertos del 20% y errores relativos demasiados altos como para utilizarlo como modelo. Se tomó la decisión de probar con una salida numérica para ver si podía subsanar el error. Los resultados fueron evidentes puesto que permitía emplear algoritmos que se adaptaban mejor al conjunto de datos.

En segundo lugar y aunque se detalle más a lo largo de las conclusiones generales, se aprecia a simple vista el poco acierto de las ecuaciones y árboles encontrados. Partimos de la base de que han sido desarrollados a partir de la inferencia de textos y documentos, aunque debería seguir un patrón establecido, no siempre tienen similitudes. Los sistemas en árbol han resultado incapaces de acertar en la predicción puesto que son más cerrados que los lineares.

Continuando con el análisis de los datos, se puede apreciar que el mejor dato se obtiene con salida numérica y eliminando algunos atributos que daban problemas. Se decide eliminar estos atributos porque tenían poco peso en la ecuación y por lo tanto apenas aportaban información o porque se obtenía un resultado contrario al esperado. Es decir, si por ejemplo un dato que en principio debía de ser positivo aparecía como negativo se eliminaba. En este conjunto de datos ha sido posible, se verá como no siempre es así.

El resultado no es demasiado apurado. Se mantiene con un error del 29%. Pero no se ha conseguido nada más preciso. En las conclusiones se darán las explicaciones sobre las posibles razones puesto que ocurre con todos los conjuntos de datos.

La ecuación obtenida en esta investigación y que será usada para el desarrollo de la aplicación es esta:

Ecuación A1
1.1753 * tamaño=reducido, grande, medio +
0.8275 * tamaño=grande, medio +
0.7888 * fijo=ecológico, local, ético, preocupación +
0.7888 * fijo=local, ético, preocupación +
0.7888 * fijo=ético, preocupación +
3.1558 * fijo=preocupación +
0.6885 * agua=ético, local, preocupación, ecológico +
0.6885 * agua=local, preocupación, ecológico +
0.6885 * agua=preocupación, ecológico +
0.7001 * agua=ecológico +
0.5109 * electricidad=preocupación +
1.598 * gestoría=ético, ecológico, preocupación, convencional +
1.598 * gestoría=ecológico, preocupación, convencional +
1.598 * gestoría=preocupación, convencional +

1.598 * gestoría=convencional +
1.2167 * prov_servicios=preocupación, ético, local, ecológico +
1.2167 * prov_servicios=ético, local, ecológico +
1.2167 * prov_servicios=local, ecológico +
1.0241 * prov_servicios=ecológico +
1.4966 * prov_material=preocupación, ético, local, ecológico +
1.4966 * prov_material=ético, local, ecológico +
1.4966 * prov_material=local, ecológico +
1.1436 * prov_material=ecológico +
3.0277 * seguros=ético, ecológico, preocupación, local +
2.4073 * seguros=ecológico, preocupación, local +
2.4073 * seguros=preocupación, local +
2.9846 * seguros=local +
1.9487 * transporte=público, ético, local, ecológico, convencional +
1.3655 * transporte=ético, local, ecológico, convencional +
1.3655 * transporte=local, ecológico, convencional +
2.0803 * transporte=ecológico, convencional +
2.4368 * transporte=convencional +
1.2763 * material_oficina=convencional, local, ético, ecológico +
1.0716 * material_oficina=local, ético, ecológico +
1.0716 * material_oficina=ético, ecológico +
1.0716 * material_oficina=ecológico +
2.2233 * material=local, ético, convencional, ecológico +
2.4589 * material=ético, convencional, ecológico +
2.4589 * material=convencional, ecológico +
2.0158 * material=ecológico +
1.9273 * ahorro_ecológico=si +
0.8275 * eficiencia_energética=si +
1.4093 * reciclaje=si +
0.9264 * proteccion_laboral=si +
1.5072 * rechazo_ético=si +
0.1268 * formación=si +
0.9685 * estudio_impacto=si +
0.9921 * encuestas_clientes=si +
2.0049 * concursos=no +
1.5072 * precios_ajustados=si +
16.4852

5.2. C1

Datos C1	Inicial	Reducida	Numérica	
Algoritmo: J48			Algoritmo: Linear Regresion	
CF: 0,01	Ac: 15.3846 % Er: 100.6029 %	Ac: 46.1538 % Er: 90.4762 %	Greedy	93.4216 %
CF:0,1	Ac: 7.6923 % Er: 104.0482 %	Ac: 23.0769 % Er: 120.3823 %	M5	89.5058 %
CF: 0,25	Ac: 7.6923 % Er: 104.7373 %	Ac: 0 % Er: 112.6984 %	NoAttrib	98.1061 %
Algoritmo: Multilayer Perceptron			Algoritmo: Multilayer Perceptron	
LR: 0,01 T: 500	Ac: 15.3846 % Er: 98.7138 %	Ac: 7.6923 % Er: 98.1357 %	LR: 0,01 T: 500	73.4409 %
LR: 0,01 T: 5000	Ac: 7.6923 % Er: 99.9819 %	Ac: 15.3846 % Er: 101.5934 %	LR: 0,01 T: 5000	73.4566 %
LR: 0,1 T: 500	Ac: 7.6923 % Er: 99.9343 %	Ac: 15.3846 % Er: 101.6022 %	LR: 0,1 T: 500	71.9916 %
LR: 0,1 T: 5000	Ac: 7.6923 % Er: 99.76 %	Ac: 15.3846 % Er: 102.37 %	LR: 0,1 T: 5000	71.8599 %
LR: 0,25 T: 500	Ac: 7.6923 % Er: 99.6695 %	Ac: 15.3846 % Er: 101.8324 %		
LR: 0,25 T: 5000	Ac: 7.6923 % Er: 99.489 %	Ac: 21.4286 % Er: 90.842 %		
Algoritmo: Random Tree				
Seed:1	Ac: 7.6923 % Er: 105.4264 %	Ac: 23.0769 % Er: 98.5714 %		
Seed: 5	Ac: 23.0769 % Er: 91.6451 %	Ac: 23.0769 % Er: 93.6508 %		
Seed: 10	Ac: 7.6923 % Er: 105.4264 %	Ac: 15.3846 % Er: 98.4127 %		

Los resultados para este atributo no son muy alentadores. Como ya se explicó se escogió C1 puesto que era el atributo más largo dentro de la tabla. La idea era probar si se adaptaba a atributos tan complicados. En A1 se ha conseguido, pero con C1 no ha sido posible.

Se ha probado de diversas formas, pero los resultados obtenidos no dan la talla. Los errores son demasiados altos y no sirven como modelo predictivo. Seguramente se deban a la inexactitud de los documentos. Ha sido muy difícil encontrar variables globales pues que en cada uno de los escritos se daba importancia a unos parámetros y a otros no. Otras veces la información no estaba clara o completa. Por lo tanto, los

algoritmos no son capaces de obtener patrones y se vuelve difícil deducir un buen modelo.

Aun así, se ha decidido mantenerlo en la memoria y en la exposición puesto que sirve como ejemplo de los límites actuales que tiene esta investigación y de que es necesario profundizar más en el método de valoración.

5.3. C3

Datos C3	Inicial	Huella Numérica	NoAttribute	Mejor
Algoritmo: Linear Regresion				
No Attribute	52.8553 %	49.6583 %	-Huella: 47.4263 %	42.6973 %
M5	63.6198 %	52.839 %	-Office: 46.9348 %	49.0914 %
Greedy	76.2259 %	78.1486 %	-Cocina: 44.9276 %	64.062 %
Algoritmo: Multilayer Perceptron			-Fbici: 43.9294 %	
LR: 0,3 T: 500	60.3402 %	64.286 %	-Formación: 42.6973 %	50.0045 %
LR: 0,1 T: 500	61.9349 %	62.9039 %	-Análisis: 45.7952 %	51.5203 %
LR: 0,01 T: 500	61.5169 %	63.6118 %		52.035 %
LR: 0,3* T: 500 S:0	60.3402 %	62.9039 %		52.035 %
LR: 0,3* T: 500 S:5	62.5711 %	62.8315 %		53.2818 %
LR: 0,3* T: 500 S:10	68.6503 %	60.0524 %		51.722 %
LR: 0,3* T: 500 S:15	65.78 %	57.365 %		51.7637 %
			Algoritmo: Tree	
			M5PTree	72.7937 %
			RandomTree	79.356

*0.1 Para el conjunto el conjunto sin la huella inicial

Como se puede apreciar para este conjunto de datos ya no se probó con la salida nominal. Teniendo en cuenta los malos resultados obtenidos se optó por comenzar directamente con la numérica. En este caso los resultados obtenidos sin haber hecho ninguna modificación no fueron del todo malos, teniendo en cuenta las tasas de error que estábamos obteniendo en general. Probablemente se deba a que es un atributo menos complejo que depende de menos variables y se puede valorar mucho mejor.

La primera mejor se obtuvo al cambiar la huella ecológica de nominal a numérica. Esta variable no estaba muy bien medida o al menos terminaba de aportar suficiente información al conjunto, aunque de primeras pareciera que sí. Por lo tanto, se decidió eliminar. Como se puede apreciar se obtuvo una mejora notable en los resultados.

El proceso que se ha seguido es el de ir eliminando atributos que no aportasen apenas información que no tuvieran sentido. Aun así, alguno no se ha podido eliminar, aunque en principio no tenga lógica. Por ejemplo, el atributo desplazamiento no termina de convencer. Primero no debería ser negativo, o al menos no en todos los casos. Aunque es cierto que si detecta que con bicicleta tiene que sumar más que con privado. Otro que tampoco termina de encajar es análisis que resta cuando tendría que sumar.

En todo caso se han intentado eliminar estos atributos, pero como los errores aumentan considerablemente se han mantenido en el modelo.

Hay que señalar que la tasa de error obtenida en el modelo final es bastante alta. Pero ha sido imposible reducirla más. Se han probado con distintos métodos, desde eliminar atributos hasta preprocesados de datos, pero en ningún caso se han obtenido mejoras.

Esta es la ecuación obtenida durante la investigación y que será empleada para el desarrollo de la aplicación:

Ecuación C3
0.6557 * tamaño=autónomos, medio, reducido +
0.2922 * tamaño=medio, reducido +
3.7414 * tamaño=reducido +
5.1537 * lugar=casa, oficina +
0.3454 * lugar=oficina +
3.4501 * vegetariano=todo, preocupación, mayoría +
3.4501 * vegetariano=preocupación, mayoría +
3.6774 * vegetariano=mayoría +
1.9425 * comida=local, ecológica, huerto +
3.6289 * comida=ecológica, huerto +
3.883 * comida=huerto +
3.6774 * formacion_aliment=si +
7.4106 * colaborativo=si +
0.6557 * llegada=bicicleta, compartido, ninguna, empresa, público, caminando +
2.5071 * llegada=compartido, ninguna, empresa, público, caminando +
2.7944 * llegada=ninguna, empresa, público, caminando +
0.3512 * llegada=empresa, público, caminando +
0.3512 * llegada=publico, caminando +
3.9141 * llegada=caminando +
0.9368* desplazamiento=privado, público, empresa, caminando, compartido, bicicleta +
-5.6625 * desplazamiento=público, empresa, caminando, compartido, bicicleta +
-1.2005 * desplazamiento=empresa, caminando, compartido, bicicleta +
-1.2005 * desplazamiento=caminando, compartido, bicicleta +
-1.2005 * desplazamiento=compartido, bicicleta +
5.3951 * desplazamiento=bicicleta +
1.0497 * ubicación=centro +
5.0867 * proximidad=si +
1.7735 * vestuario=si +
2.1662 * virtuales=si +
2.02 * fpublico=si +
3.776 * faparcamiento=si +
4.063 * ejemplaridad=si +
2.5802 * incentivos=si +
1.5045 * material=si +

11.2169	* practicas=si +
2.0676	* directivos=empresa, ninguna, público, compartido, caminando, bicicleta +
2.0676	* directivos=ninguna, público, compartido, caminando, bicicleta +
2.0676	* directivos=público, compartido, caminando, bicicleta +
2.2637	* directivos=compartido, caminando, bicicleta +
2.7634	* directivos=caminando, bicicleta +
2.4267	* directivos=bicicleta +
-0.96	* análisis=si +
2.0668	

6. Análisis del sistema

En este punto se recoge el análisis de uso del sistema final. En este prototipo final quedan reflejadas todas las conclusiones que se han obtenido en la fase de investigación. Se ha separado en dos partes la documentación puesto que son fases distintas y corresponden a ideas distintas dentro del proyecto.

Durante este apartado haremos una descripción del sistema que se ha desarrollado. Más adelante se explicará técnicamente en que consiste y como se ha llevado a cabo. La aplicación programada durante este proyecto es una demo que servirá para explicar gráficamente los resultados de la investigación

6.1. Casos de uso

Los casos de uso se han decidido elaborar en formato expandido para poderlo detallados de la mejor forma posible. Estas son las características que se podrán apreciar en cada uno de ellos. Se definen ahora para facilitar la lectura de los mismos.

- **Identificador:** únicos para cada caso de uso. Se identificarán con un acrónimo de la utilidad que representan (máximo 4 letras) seguido de un número en caso de que haya varios con la misma. Por ejemplo, si un caso es pedir información el identificador será CU-INFO-01.
- **Actores:** los actores que realizan el caso de uso en concreto. En nuestro caso tendremos SISTEMA y USUARIO.
- **Propósito:** el objetivo que se busca con el caso de uso en concreto.
- **Visión General:** Resumen o explicación del caso de uso.
- **Tipo:** importancia del caso de uso. Primario, secundario u opcional.
- **Referencias:** Relaciones con otros casos de uso.
- **Curso Típico de los Eventos:** Pequeño esquema del desarrollo de acciones que implica el caso de uso.
- **Cursos alternativos:** Alternativas a la consecución natural de los hechos expresados en el caso de uso.

Identificador	CU-FORM-01	
Actores	Usuario y Sistema	
Propósito	Rellenar el Formulario	
Visión General	El usuario siente la necesidad de obtener la puntuación y comienza su proceso para ello. Consiste en rellenar el formulario que contiene todas las variables	
Tipo	Primario	
Referencias		
Curso típico de los eventos	1. El usuario inicia la aplicación.	
		2. La aplicación le muestra el

		formulario por pantalla.
	3. El usuario rellena todos los campos del Formulario	

Identificador	CU-FORM-02	
Actores	Usuario Y Sistema	
Propósito	Pedir información de los campos	
Visión General	En algún atributo puede surgir la duda de que significa o que opciones puede elegir. Para ello el usuario coloca el ratón sobre el icono de información y recibe la información necesaria	
Tipo	Secundario	
Referencias		
Curso Típico de los eventos	1. El usuario coloca el ratón en el icono información cercano al campo del formulario	
		2. El sistema muestra un cuadro informativo con todo lo necesario para comprender el atributo a rellenar.

Identificador	CU-PUNT-01	
Actores	Usuario y Sistema	
Propósito	Pedir la puntuación final	
Visión General	El usuario una vez a rellenado los campos necesarios le pide al sistema que le muestra la calificación obtenida	
Tipo	Primario	
Referencias		
Curso Típico de eventos	1. El usuario presiona el botón de calcular puntuación.	
		2. El sistema realiza los cálculos pertinentes.
		3. El sistema muestra por pantalla un anexo

		con la puntuación obtenida.
Curso Alternativo		2. El sistema calcula un error porque el usuario ha dejado algún campo sin rellenar.

Identificador	CU-PUNT-02	
Actores	Sistema	
Propósito	Calculo de la puntuación	
Visión General	El sistema recoge las indicaciones de los campos del formulario las procesa y genera la puntuación final	
Tipo	Primario	
Referencias		
Curso Típico de eventos		1. El sistema obtiene el valor de los atributos del formulario.
		2. El sistema introduce los valores en la fórmula de predicción.
		3. El sistema obtiene la puntuación final tras haber procesado los datos.

Identificador	CU-EXCL	
Actores	Usuario y Sistema	
Propósito	Generar un excell de puntuaciones	
Visión General	Una vez rellenados todos los campos de los 17 atributos y calculadas las puntuaciones el usuario puede solicitar descargar un archivo con todas ellas en formato .xls	
Tipo	Opcional	
Referencias		
Curso Típico de eventos	1. El usuario rellena todos los campos.	

	2.El usuario solicita la puntuación	
		3.El sistema genera las puntuaciones requeridas
		4. El sistema muestra el botón de generar la matriz en formato excell.
	5.El usuario pulsa el botón para generar el excell	
		6. El sistema genera un archivo excell con la matriz y los atributos calculados
	7. El usuario guarda el archivo en su sistema.	
Curso Alternativo		3. El sistema arroja un error porque el usuario se ha dejado algún campo en blanco
	4. El usuario rellena el campo faltante.	

6.2. Requisitos

Este punto se centra en describir todos aquellos requisitos necesarios para que la aplicación en desarrollo cumpla con la función exigida. Los requisitos han ido variando a lo largo del desarrollo conforme se encontraban nuevos o alguno se veía incompleto.

El sistema de clasificación de requisitos sigue los estándares. Tendremos en primer lugar dos tipos principales: los funcionales y los no funcionales. Dentro de esto últimos de sistema, de usabilidad y de interfaz. Igual que sucedió en el apartado de casos de uso, se adjunta una guía con las explicaciones pertinentes sobre las características de cada requisito.

- Identificador: RX-Y-NN. Siendo RX las iniciales de Requisito Funcional o No funcional. La Y indica que subtipo es el requisito (sistema, usabilidad e interfaz). Por último, la NN indica el número de requisito.
- Necesidad: indica que importancia requiere dentro del sistema
- Prioridad: con la que tomará en cuenta el requisito
- Claridad: grado de ambigüedad del requisito
- Verificabilidad: facilidad de comprobación.
- Descripción. Una pequeña definición o explicación del requisito.

6.2.1. Requisitos Funcionales

RF-01	
Necesidad: media	Prioridad: media
Claridad: Alta	Verificabilidad: Alta
Descripción	El sistema tendrá que mostrar información sobre los atributos de cada uno de los indicadores cuando se requiera. Cuando se requiera con el rato encima o con un click.

RX-Y-02	
Necesidad: Alta	Prioridad: Alta
Claridad: Alta	Verificabilidad: Media
Descripción	El sistema cuando se pulse el botón de calcular ha de mostrar por pantalla la valoración en función de los datos introducidos.

RX-Y-03	
Necesidad: Baja	Prioridad: Baja
Claridad: Media	Verificabilidad: Alta
Descripción	Generar archivo xls con los 17 indicadores.

6.2.2. Requisitos No funcionales

6.2.2.1. Requisitos del Sistema

RNF-S-01	
Necesidad: Media	Prioridad: Media
Claridad: Alta	Verificabilidad: Media
Descripción	La aplicación web ha de ser compatible con todos los navegadores del mercado para facilitar su uso.

RNF-S-02	
Necesidad: Alta	Prioridad: Alta
Claridad: Alta	Verificabilidad: alta

Descripción	El lenguaje del sistema ha de ser javascript, combinado con html y css.
--------------------	---

RNF-S-03	
Necesidad: Alta	Prioridad: media
Claridad: Alta	Verificabilidad: alta
Descripción	Las librerías que se emplean para parte de la interfaz han de estar desarrolladas en JQuery

6.2.2.2. Requisitos de Usabilidad

RNF-U-01	
Necesidad: Media	Prioridad: Media
Claridad: Baja	Verificabilidad: baja
Descripción	Toda la interfaz del sistema ha de ser simple y de fácil uso. Para ello todo tiene que ser intuitivo de fácil comprensión

RNF-U-02	
Necesidad: Alta	Prioridad: Media
Claridad: Media	Verificabilidad: Media
Descripción	Los atributos de cada indicador han de ser perfectamente entendibles y no dar lugar a equivocación para facilitar en análisis de la persona que rellena el formulario.

RNF-U-03	
Necesidad: Baja	Prioridad: Baja
Claridad: Alta	Verificabilidad: Alta
Descripción	El formulario ha de tener desplegables para que rellenen los campos. De esta forma se evitan equivocaciones.

RNF-U-04	
Necesidad: Media	Prioridad: Media
Claridad: Alta	Verificabilidad: Media
Descripción	El sistema ha de adaptarse a cada pantalla donde vaya a ser ejecutado. Desde móviles a ordenadores de sobremesa.

6.2.2.3. Requisitos de Interfaz

RNF-I-01	
Necesidad: Alta	Prioridad: Alta

Claridad: Alta	Verificabilidad: Alta
Descripción	La interfaz ha de disponer de un formulario para rellenar todos los atributos de cada indicador.

RNF-I-02	
Necesidad: Media	Prioridad: Media
Claridad: Alta	Verificabilidad: Alta
Descripción	En cada uno de los indicadores se ha de disponer de un cuadro de ayuda para facilitar su entendimiento

RNF-I-03	
Necesidad: Alta	Prioridad: Media
Claridad: Alta	Verificabilidad: Alta
Descripción	Cada uno de los formularios de los indicadores ha de disponer de un botón calcular que muestre por pantalla la información obtenida.

RNF-I-04	
Necesidad: Media	Prioridad: Baja
Claridad: alta	Verificabilidad: Alta
Descripción	Después de todos los indicadores ha de aparecer un botón que cree un archivo Excel que pueda ser descargado por el usuario. En este Excel estarán almacenadas todas las puntuaciones.

RNF-I-05	
Necesidad: Alta	Prioridad: Alta
Claridad: Alta	Verificabilidad: Alta
Descripción	Han de existir 17 formularios en la interfaz, uno por cada indicador.

RNF-I-06	
Necesidad: Alta	Prioridad: Alta
Claridad: Alta	Verificabilidad: Alta
Descripción	Cada uno de los formularios de los indicadores ha de disponer de todos los atributos necesario para su valoración.

7. Diseño la Demo

7.1. Entorno de Desarrollo

El desarrollo del proyecto se ha llevado a cabo en un entorno de desarrollo estándar. En ambos ordenadores se trabajó sobre Windows 10 Pro. Al principio se comenzó a desarrollar en una máquina virtual con Linux, pero puesto que no había necesidad de seguir en esa plataforma se migraron los archivos al sistema principal y se trabajó desde ahí.

Para la programación se ha empleado Notepad++. Este programa es un editor de código gratuito que soporta múltiples lenguajes de programación. Está programado en C++ y emplea una Api de Windows para ejecutarse. Además, en un software que intenta reducir las emisiones de carbono puesto que reduce el consumo de energía del ordenador al requerir menos uso de la CPU.

El lenguaje de programación escogido ha sido HTML para la estructura de la web, CSS para los diseños y algunas animaciones y JavaScript para el evaluador de empresas. Para la interfaz gráfica se han implementado soluciones bootstrap para mejorar el rendimiento, así como la imagen de la página.

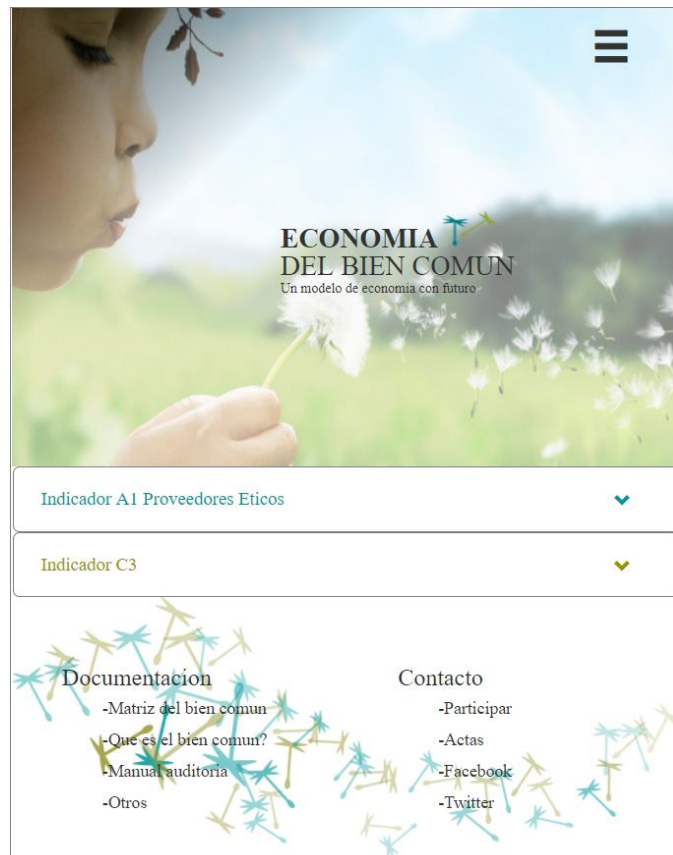
7.2. Clasificador

El sistema se compone de dos partes diferentes, la interfaz y las funciones de evaluación. Por lo tanto, se dedicará un apartado a cada una. En cada uno se explicará brevemente el código desarrollado y se anexarán capturas para que resulten más fácil la comprensión.

El código se compone de varios archivos:

- Formulario.html: se define la estructura de la página y el código de valoración de empresa.
- Formulario.css: contiene los detalles visuales de la página
- Formulario.js: contiene algunas animaciones de la página.

La demo básicamente consiste en dos formularios donde rellenar los distintos atributos de los dos indicadores y obtener la valoración final. De esta forma se puede apreciar lo que se ha conseguido en las fases previas del proyecto. En la aplicación final tendría los 17 indicadores.



A1 Gestión ética de los suministros >

Tamaño de la empresa	...
Contrato Teléfono Fijo	...
Contrato Agua	...
Contrato Electrico	...
Gestaria	...
Proveedores de servicios	...
Proveedores de materiales	...
Seguro contratado	...
Transporte empleado	...
Tipo de Material en la Oficina	...
Tipo de material usado en la producción	...
Empleo de técnicas de ahorro ecológico	...
Empleo de técnicas para la eficiencia energética	...
Medidas de Reciclaje	...
Protección laboral en proveedores	...
Rechazo ético de proveedores	...
Cursos de Formación	...
Estudio de impacto de proveedores	...
Encuestas a proveedores	...
Concursos en la contratación	...
Precios ajustados al los proveedores	...
Total	0

Calcular A1

C3 Promoción del comportamiento ecológico v

A1 Gestión ética de los suministros v

C3 Promoción del comportamiento ecológico >

Numero de empleados	...
Lugar habitual de comidas	...
Iniciativas en la dieta vegetariana	...
Tipo de productos consumidos habitualmente	...
Formación Alimentaria	...
Proyecto colaborativo para la alimentación	...
Modo de llegada al puesto	...
Modo de transporte en el puesto	...
Ubicación de la empresa	...
Personal en las proximidades	...
Existencia de vestuarios	...
Empleo de videoconferencias	...
Facilidad acceso en T.público	...
Facilidad de Aparcamiento	...
Ejemplaridad ante otras empresas	...
Incentivos al comportamiento eco	...
Material para la formación	...
Prácticas ecológicas	...
Transporte usado por la dirección	...
Estudio sobre la contaminación de la empresa	...
Total	0

Calcular C3

7.2.1. Interfaz

La interfaz se ha intentado desarrollar bajo los criterios gráficos de la EBC. Para ello se ha seguido el manual de diseño del que disponen. Esto ha servido para elegir la paleta de colores a emplear y alguna de las imágenes.

El código del formulario se ha estructurado en tantos <inputs> como atributos tuviera el indicador. A cada uno se le da un id propio que será necesario para el código de evaluación. De esta forma podemos acceder a cualquiera de estos atributos en cualquier momento. Esto también es útil a futuro por si se quiere desarrollar alguna otra utilidad con estos formularios.

```
<div id="formulario_a" class="input-group input-group-lg">
  <span class="input-group-addon" id="sizing-addon1" style="background-color: rgb(0,153,153); color:white;">Tamano de la empresa</span>
  <input id="Tamano" list="formaTamano" type="text" class="form-control" placeholder="..." aria-describedby="sizing-addon1" >
  <datalist id="formaTamano">
    <option value="Autonomo">
    <option value="2-10">
    <option value="11-50">
    <option value="+50">
  </datalist>
</div>
```

Se ha decidido modelar la interfaz como si fuera una página completa para redondear un poco la imagen visual. Para ello se ha implementado un pequeño menu que contiene los enlaces a las otras secciones de la aplicación y un glosario en la parte inferior donde se recogen los enlaces a la documentación así como la forma de contacto. Así cuando se quiera continuar con el trabajo se puede implementar de forma de web con distintas páginas a parte del clasificador.

```
<ul class="dropdown-menu" aria-labelledby="dropdownMenu1">
  <li><a href="#">Que es la EBC?</a></li>
  <li><a href="#">Clasificador </a></li>
  <li><a href="#">Glosario Empresas</a></li>
  <li><a href="#">Documentacion</a></li>
  <li role="separator" class="divider"></li>
  <li><a href="#">Contacto</a></li>
</ul>

<div class="contacto">
  <div class="div_contacto">
    <h1 class="contacto_titulo">Documentacion</h1>
    <p class="contacto_texto">-Matriz del bien comun </p>
    <p class="contacto_texto">-Que es el bien comun?</p>
    <p class="contacto_texto">-Manual auditoria </p>
    <p class="contacto_texto">-Otros</p>
  </div>
  <div class="div_contacto">
    <h1 class="contacto_titulo">Contacto</h1>
    <p class="contacto_texto">-Participar </p>
    <p class="contacto_texto">-Actas</p>
    <p class="contacto_texto">-Facebook </p>
    <p class="contacto_texto">-Twitter</p>
  </div>
```


7.2.2. Evaluador

```

/**
 * Funcion que se ejecuta cada vez que hace click en el boton de valorar
 * Reconoce los valores del formulario y calcula el valor en función de estos
 * En concreto calcula el valor del Atributo A1
 */
function sumarA1()
{
    var atribA1= 16.4852;
    var tamano = document.getElementById("Tamano").value;
    if(tamano=="Autonomo")
        atribA1= atribA1 + 0;
    if(tamano=="2-10")
        atribA1= atribA1 + 1.1753;
    if(tamano=="11-50")
        atribA1= atribA1 + 2.0028;
    if(tamano==">50")
        atribA1= atribA1 + 2.0028;
    if(tamano!=">50" && tamano!="11-50" && tamano!="2-10" && tamano!="Autonomo" )
    {
        // marcamos como erroneo
        document.getElementById("Tamano").style.borderColor="#f00";
        return 0;
    }

    var fijo = document.getElementById("Telefono").value;
    if(fijo=="Convencional")
        atribA1= atribA1 + 0;
    if(fijo=="Ecologico")
        atribA1= atribA1 + 0.7888;
    if(fijo=="Entidad local")
        atribA1= atribA1 + 0.7888 + 0.7888;
    if(fijo=="Etico")
        atribA1= atribA1 + 0.7888 + 0.7888 + 0.7888;
    if(fijo=="Alguna Preocupacion")
        atribA1= atribA1 + 0.7888 + 0.7888 + 0.7888 + 3.1558;

    if(fijo!="Convencional" && fijo!="Alguna Preocupacion" && fijo!="Entidad local" && fijo!="Etico" && fijo!="Ecologico" )
    {
        // marcamos como erroneo
        document.getElementById("Telefono").style.borderColor="#f00";
        return 0;
    }

    document.getElementById("valorC3").value= Math.round(atribC3);
}

```

Para mejorar la comprensión en la explicación del código adjunto un resumen del evaluador solo con atributos. Esta función esta implementando los botones valorar de cada atributo. En primer lugar se da valor al atributo en función del termino independiente de la ecuación.

En el caso de A1 16,4882. Después y para cada atributo se obtiene el valor del formulario que le corresponde. En este caso para el tamaño de la empresa y el tipo de contrato para el teléfono fijo.

Una vez obtenidos los valores de los campos se implementan condicionales específicos para cada uno de ellos. De tal forma que suma al valor final en función del de cada atributo. Se implementa también la función de error en caso de que algun campo se deje sin rellenar o se le de un valor erroneo. En caso de error se marcará el recuadro en rojo y realizará la suma.

Por último, se redondea el valor de A1 y se muestra por pantalla.

8. Plan Económico

8.1. Planificación del Proyecto

El proyecto tiene una duración total de 4 y tres semanas. Es decir, 95 días laborables. Cada día de dedicación ha supuesto una media de 6 horas, por lo tanto, han sido un total de 570 horas. Como ya se expuso en los primeros apartados, el proyecto se ha estructurado en distintas fases de desarrollo que abarcan desde la investigación hasta la redacción de la memoria. A continuación, se detallan los momentos de desarrollo y su duración determinada de la manera más exacta posible.

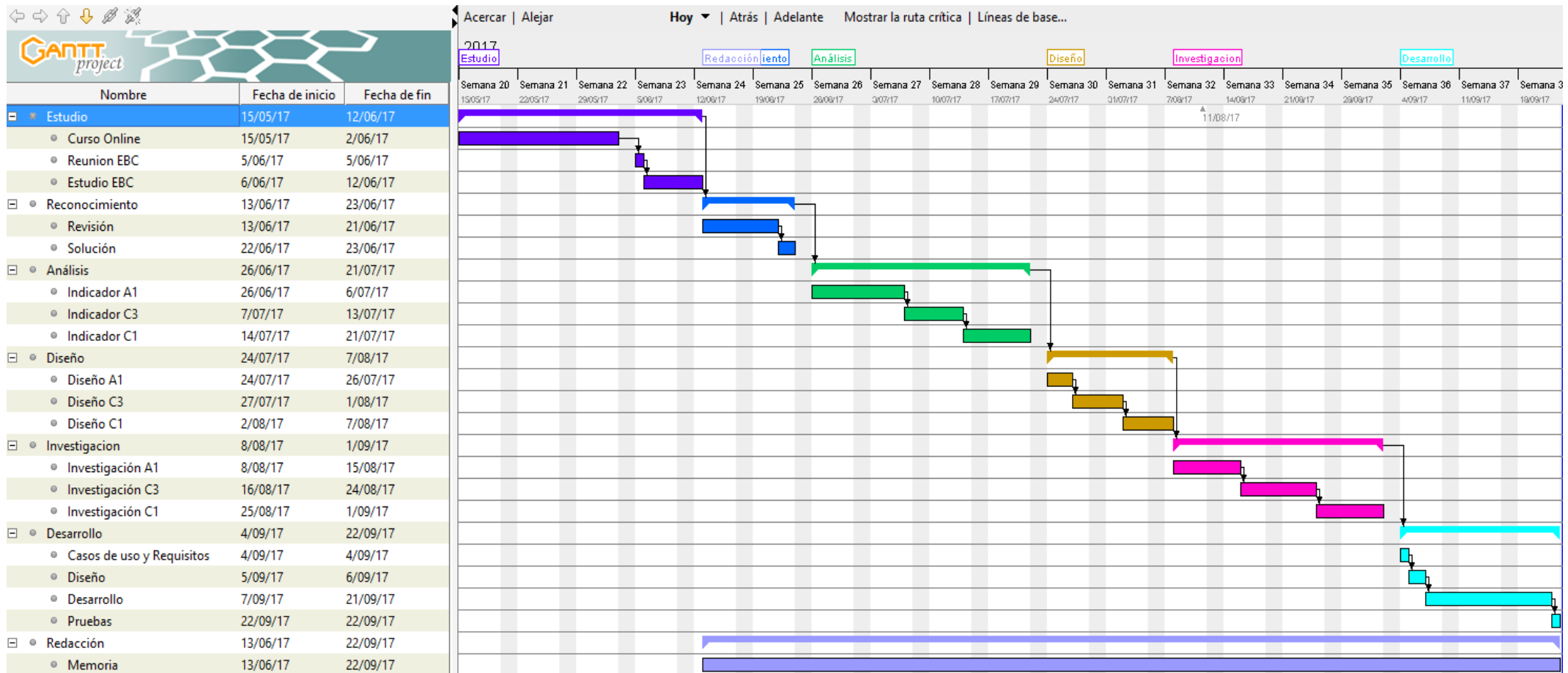
1. **Estudio:** Esta primera fase consistió en estudiar toda la documentación que disponía de la economía del bien común y el curso de la plataforma coursera sobre aprendizaje automático. En esta fase básicamente se trató de contextualizar el objeto del proyecto y contar con el conocimiento necesario para poder reconocer el problema y tratar de darle solución.
Fue una de las fases más largas puesto que se tuvo que recabar información de las fuentes necesarias y cursar el curso online (aunque no en su totalidad). La duración de esta fase fue 21 días. En este caso la dedicación fue completa y no se fue redactando la memoria mientras tanto. Por ello el momento de desarrollo que más tiempo ha durado.
2. **Reconocimiento del problema.** Una vez recabada toda la información se pasó a encontrar la forma en la que se podría aplicar el aprendizaje automático en el campo de la economía del bien común. Para ello se revisó la técnica con la que realizaban la evaluación y se dio con el problema de fondo. Esta fase tuvo una duración de 9 y ya se empezó a redactar la memoria al mismo tiempo.
3. **Análisis.** Esta fase fue una de las más pesadas de llevar a cabo. Se tuvo que analizar uno por uno los documentos de evaluación de la economía del bien común. Para cada uno de los atributos se tuvo que buscar un modelo que pudiera ser efectivo para todos los casos. Este proceso incluye ir comparando los apartados entre sí y con el modelo teórico proporcionado por la EBC. Además, había que generar los conjuntos de datos iniciales lo cual también lleva un tiempo importante. Fue un trabajo lento y pormenorizado que llevo un tiempo considerable, unos 20 días.
4. **Diseño.** Una vez recabada toda la información y obtenidos los modelos iniciales, había que pensar en las pruebas que se iban a realizar. Había que pensar en que algoritmos iban a ser útiles y cuales se iban a implementar en cada caso. También se pensó en variaciones de los conjuntos iniciales de datos con las que investigar. Este proceso no fue excesivamente largo y duro 11 días

5. **Investigación** Una vez dejado todo definido el trabajo se centró en la investigación con los datos de los que se disponía a través de la plataforma weka. Las pruebas se realizar tal y como se había planificado, pero hubo que añadir más conforme se iban encontrando mejores conjuntos de datos. Después de esta investigación se obtuvieron las ecuaciones de predicción para cada uno de los modelos que se habían descrito. Tuvo una duración de 19 días puesto que había que realizar numerosas pruebas e ir analizando la salida de cada uno de ellas.
6. **Desarrollo**. Esta fue la última fase del proyecto. Con las ecuaciones de las que disponía había que elaborar una aplicación básica que mostrará el sistema de valoración automático que se había teorizado. Esta fase abarcad desde el pensamiento de los casos de uso y los requisitos hasta la programación final del formulario web y el código que calcularía el valor de cada atributo. Conto con un total de 15 días.
7. **Redacción**. Esta fase ha sido transversal, se ha ido redactando la memoria conforma se avanzaba en el proyecto. Se ha trabajado una media de una hora al día. Esto significa que, aunque había días que no se dedicaba tiempo, otros se dedicaban 2 o 3 horas.

En esta tabla se resumen las duraciones de las fases para que quede claro la distribución del trabajo.

Fase	Días	Horas/Día	Horas
Estudio	21	6	126
Reconocimiento	9	5	45
Análisis	20	5	100
Diseño	11	5	55
Investigación	19	5	95
Desarrollo	15	5	75
Redacción	74	1	74
Total	95	6	570

8.2. Diagrama de Gantt



8.3. Costes

Durante este punto se explica el coste total del proyecto. Se desglosará el total en las diferentes partidas para clarificar. Puesto que este proyecto no es más que el inicio de algo más grande, se ha decidido calcular los costes de tan solo de esta primera muestra. Los costes general

8.3.1. Personal

Puesto que solo ha habido un trabajador que se ha dedicado a varias tareas se muestra el coste en función de las tareas que ha desempeñado. En la parte de análisis se engloba toda la investigación dedicada a obtener las ecuaciones de predicción del sistema. Esta primera parte es la que más horas lleva debido a que es donde reside el grueso del sistema. El papel de programador se centra en desarrollar la aplicación que sirva como evaluador de empresas. Esta parte es más sencilla y requiere menos hora. Por último, se necesita de una persona que realiza las pruebas sobre el sistema desarrollado. Cada trabajador hace al día 6 horas, máximo 120 al mes

Nombre	Puesto	Dedicación	Coste Hora	Coste
Alejandro	Analista/investigado	421 horas	13,02€/hora	5.480,90
Alejandro	Programador web	60 horas	12,50€/hora	750,00
Alejandro	Tester	15 horas	10,93 €/hora	163,95
Total				6.394,85

Todos los salarios han sido calculados en base a referencias encontradas en la web desde portales de empleo.⁸

8.3.2. Hardware

En este punto se detalla el coste del hardware utilizado durante el proyecto. Todos los costes están calculados según su precio real de venta sin IVA. Para calcular el coste del equipo se emplea la formula

$$A/B * C * D$$

⁸<http://espana.jobtonic.es/salary/26526/16119.html>
<http://espana.jobtonic.es/salary/26526/84380.html?currency=EUR>
<http://espana.jobtonic.es/salary/26526/16272.html>

Siendo:

- A: nº de meses desde la fecha de facturación en que el equipo es utilizado
- B: periodo de depreciación (60 meses)
- C: coste del equipo (sin IVA)
- D: % del uso que se dedica al proyecto (habitualmente 100%)

Hardware	Cantidad	Coste	Utilización	Coste Imputable
Ordenador Sobremesa	1	835 €	4 meses	55,67
Ordenador Portátil: Asus	1	650 €	1 mes	10,83
Total	2	1485 €	5 meses	66,50€

8.3.3. Software

Todo el software que se ha utilizado ha sido gratuito o con licencias educativas. Por lo tanto, se imputan solo los gastos de las licencias de estas licencias para mostrar el coste real. El coste se obtiene del precio de venta de las páginas oficiales.

Software	Coste
Weka	0 €
GanttProject	0 €
Windows 10	130 €
Office 2016	70 €
Notepad++	0 €
Total	200 €


8.3.4. Total

En esta tabla se resume el coste total del proyecto.

Asunto	Coste
Costes Persona	6395,73 €
Costes Hardware	66,49 €
Costes Software	200 €
Costes indirectos (20%)	1.332
Total	7.994 €

Al final el coste total del proyecto asciende a 6.662,22 euros repartidos en las distintas partidas de gasto.

8.4 Presupuesto Detallado

 UNIVERSIDAD CARLOS III DE MADRID Escuela Politécnica Superior							
PRESUPUESTO DE PROYECTO							
1.- Autor:							
Alejandro Layos							
2.- Departamento:							
Informática							
3.- Descripción del Proyecto:							
- Título	CLASIFICADOR DE EMPRESAS BAJO CRITERIOS DE LA EBC						
- Duración (meses)	5						
Tasa de costes Indirectos:	20%						
4.- Presupuesto total del Proyecto (valores en Euros):							
Euros							
5.- Desglose presupuestario (costes directos)							
PERSONAL							
Apellidos y nombre	N.I.F. (no rellenar - solo a título informativo)	Categoría	Dedicación (hombres mes) ^{a)}	Coste hombre mes	Coste (Euro)	Firma de conformidad	
Alejandro Layos		Analista	3,508	1.562,40	5.480,90		
Alejandro Layos		Programador Web	0,5	1.500,00	750,00		
Alejandro Layos		Tester	0,125	1.311,60	163,95		
					0,00		
					0,00		
Hombres mes 4,133				Total	6.394,85		
^{a)} 1 Hombre mes máximo 120 al mes							
EQUIPOS							
Descripción	Coste (Euro)	% Uso dedicado proyecto	Dedicación (meses)	Periodo de depreciación	Coste imputable ^{d)}		
Ordenador Sobremesa	835,00	100	4	60	55,67		
Ordenador Portátil: Asus	650,00	100	1	60	10,83		
		100		60	0,00		
		100		60	0,00		
		100		60	0,00		
					0,00		
					Total	66,50	
^{d)} Fórmula de cálculo de la Amortización:							
$\frac{A}{B} \times C \times D$		A = nº de meses desde la fecha de facturación en que el equipo es utilizado B = periodo de depreciación (60 meses) C = coste del equipo (sin IVA) D = % del uso que se dedica al proyecto (habitualmente 100%)					
SUBCONTRATACIÓN DE TAREAS							
Descripción	Empresa	Coste imputable					
		Total	0,00				
OTROS COSTES DIRECTOS DEL PROYECTO^{e)}							
Descripción	Empresa	Costes imputable					
Software		200,00					
		Total	200,00				
^{e)} Este capítulo de gastos incluye todos los gastos no contemplados en los conceptos anteriores, por ejemplo: fungible, viajes y dietas,							
6.- Resumen de costes							
Presupuesto Costes Totales	Presupuesto Costes Totales						
Personal	6.395						
Amortización	67						
Subcontratación de tareas	0						
Costes de funcionamiento	200						
Costes Indirectos	1.332						
Total	7.994						

9.Conclusiones

A lo largo de este punto intentaré resumir el alcance de los objetivos marcados al principio del documento, concluir con todo el proceso de investigación que he llevado a cabo y aportar mi visión personal.

En primer lugar, comenzaré explicando el resultado en función de los objetivos que me propuse al iniciar este trabajar. Al principio de este proyecto planteé 3 objetivos a cumplimentar a lo largo de la investigación y el desarrollo. El primero de ellos el de la aplicación de algoritmos de Machine learning al proceso de evaluación. Es objetivo es el que más tiempo ha requerido a lo largo del proyecto puesto que abarca toda la investigación y el modelado de datos. Ha supuesto un trabajo manual y tuvo que hacerse de manera lenta y minuciosa. Obtener los atributos para los conjuntos de datos a través de los informes fue laborioso, pero al final se pudo conseguir. Una vez obtenidos los conjuntos se pudo aplicar sin mayores problemas los algoritmos de aprendizaje automático. No todos encajaban y, como es de esperar, no todos arrojaron buenos resultados, pero sí que se pudo demostrar que su aplicación era viable y habría una nueva vía de desarrollo.

El siguiente objetivo a repasar es el de un prototipado de un evaluador automático. Los resultados se pueden apreciar en el formulario. Aunque tan solo es una muestra y no es un diseño final, si se ve para ver su utilidad. A través de este formulario se consigue evaluar de forma automática dos de los 3 atributos estudiados. Por último, hablare del objetivo general de este proyecto, automatizar el sistema de evaluación de la EBC. Este objetivo se ha conseguido parcialmente. Está claro que no se ha conseguido obtener un automatizador completo del proceso, pero si se han dado los primeros pasos para su desarrollo final. A lo largo del trabajo se aportan las pautas que sirven para estudiar cada atributo y obtener su conjunto de datos. Si se continuase con este estudio se podría llegar a obtener un modelo final que sirviera para sustituir en un futuro al proceso manual de evaluación.

En cuanto a mis apreciaciones acerca de la aplicación de esta tecnología creo que usar aprendizaje automático en el proceso de evaluación es una buena idea a futuro para poder solucionar los problemas que describí al principio del documento. Con las ecuaciones se reduce el tiempo necesario para poder evaluar, se elimina la necesidad de una formación previa para la realización de los documentos y se prescinde de hacer que las personas tengan que desplazarse de manera asidua a las empresas.

Es una tecnología que se adapta a este tipo de iniciativas en proceso de crecimiento y de creación constante. Reporta mucha utilidad porque con los documentos elaborados se pueden generar ecuaciones que se ajusten a la realidad y que puedan ir cambiando en caso de que se cambien criterios y o valoraciones. En caso de que algún criterio perdiese peso, bastaría con volver a ajustar los datos y volver a

correr los logaritmos sobre el conjunto. O incluso con que modificaran las ecuaciones bajando el peso específico a la variable en cuestión también valdría. Teniendo en cuenta que están cambiando algunos parámetros cada cierto tiempo resulta importante esta adaptabilidad.

Como ya he dicho este trabajo es una primera aproximación hay 17 indicadores y solo he trabajado en torno a 3 de ellos. Manejar los 17 hubiera supuesto un tiempo del no que disponía. Sigue siendo una iniciación y los resultados no son definitivos y totalmente concluyentes. Como se puede apreciar en la precisión de las ecuaciones aún falta precisión a la hora de predecir, los errores son altos. En mi humilde opinión y con la experiencia de estos meses creo que se debe a varias circunstancias. Comentare dos problemas primero que se refieren a la parte teórica o de la elaboración de sistema. El primero es el error humano. Parte de mi trabajo ha consistido en sistematizar en forma de variables unos documentos basándome en los manuales de la EBC y en la comparación entre ellos. Este proceso está sujeto a errores. Puedo haber interpretado variables que no influyen o puedo haber creado atributos que no tenían consistencia que simplemente no se adaptaban a la realidad. En segundo lugar, creo que no hay coherencia o escalabilidad en los documentos. Es decir, creo que algunas empresas se sobrevaloran y en otros casos se infravaloran. Durante el estudio de las variables percibí que empresas que a priori no cumple excesivamente con los valores o que no tiene medidas en un punto en concreto tiene puntuaciones infladas. En cambio, otras parecen tener la misma puntuación con muchas mejores iniciativas. Sería bueno revisar o por lo menos asegurar que se cumple el criterio general que es establece en los manuales. Como ya indiqué esta subjetividad a la hora de crear la herramienta teórica de valoración redunda en los errores obtenidos. Si el método de evaluación fuera algo más preciso se lograría obtener calificaciones más ajustadas a la realidad, que es básicamente lo que buscaba con este proyecto y que en cierto punto parece funcionar.

Continuando con los problemas que he visto en la realización del trabajo, comentaré ahora dos problemas que son algo más técnicos. El primero es la falta de ejemplos de entrenamiento. En total he tenido 16 documentos de 16 empresas distintas, todas de tamaños dispares. Además 2 de ellas no tenían la valoración en forma de número. Es decir que había 14 ejemplos. Esto ha sido uno de los principales problemas a la hora de establecer conclusiones. Con un conjunto tan reducido es complicado crear una ecuación en la que podamos asegurar que se va obtener la puntuación de forma precisa para cualquier caso. En segundo lugar y relacionado con el primero aparece el problema de overfitting. Al tener tan pocos ejemplos de entrenamiento puedo haber conseguido una ecuación que se adecue a los ejemplos previos pero que no sirva para nuevas empresas. Pero, pese a que son problemas relativamente importantes tiene fácil solución, conforme avance el tiempo y se disponga de más documentos se puede hacer pruebas más detalladas con varios conjuntos de entrenamiento y obtener mejores ecuaciones.

En el terreno personal pienso que este trabajo me ha aportado mucho. He aprendido a trabajar sobre el terreno. Mantener reuniones con la organización,

comunicarles avances, pedir documentaciones o incluso consultarles dudas a la vez que realizaba el trabajo ha sido muy buena experiencia. De cara al mundo laboral poder trabajar con clientes y saber gestiones de forma correcta el trato con ellos es fundamental en el desarrollo de proyectos. No solo ha sido el trato con cliente, también el uso de un proyecto real como es el desarrollado por la EBC, todas las empresas que colaboran, el hecho de emplear tecnologías al alcance de cualquiera o el desempeño técnico del trabajo creo que le confieren un sentido amplio al trabajo que detallo. Además, ha sido muy enriquecedor poder aplicar o investigar la aplicación de una tecnología en un terreno donde todavía no se había aplicado.

Me ha servido para aprender, tanto sobre la economía del bien común como sobre aprendizaje automático. Durante la realización del trabajo estuve estudiando un curso online de Machine learning que me vino bien para enfocarlo mucho más a la gestión de datos que inteligencia artificial como he aprendido en la carrera. Además, le hecho de poder aplicarlo a un caso real ha sido muy satisfactorio puesto que he podido ver sus posibilidades y su utilidad real. Además, he tenido que trabajar mucho sobre la EBC y leerme y estudiarme sus materiales y comprenderlos para después poder aplicarlos en mi desarrollo.

Es complicado decir si esta corriente económica va a llegar lejos o si se quedará en una pequeña iniciativa ciudadana. Mi idea principal a la hora de encontrar el objetivo de mi trabajo de fin de carrera de informática era la de crear un proyecto que sirviera o al menor poder apoyar a alguna organización que no tendría los medios por si misma de desarrollar este tipo de proyectos.

10.Trabajo futuro

Abriría varias líneas de trabajo futuro en las que centrar a la próxima persona que vaya a trabajar en este tema:

- **Estudiar los 17 Indicadores.** terminaría de realizar el estudio para los indicadores que falta y completaría la tabla de evaluación. De esta forma se podrá disponer de un completo sistema para puntuar a las empresas de forma automática y a través de formularios para rellenar.
- **Recaudar más evaluaciones.** puesto que es un trabajo a futuro, deduzco que más adelante habrá muchas más evaluaciones y se podrá realizar un trabajo más completo. Habría que analizarlas e introducirlas en los conjuntos de datos y trabajar sobre ellas como he intentado hacer durante este estudio inicial.
- **Buscar un equipo de trabajo.** Trabajar en los 17 indicadores de forma concienzuda para una sola persona es demasiado trabajo. Si se quiere encontrar buenos modelos de trabajo lo mejor es dejar que un equipo de varias personas se encargue de repartírselos. Con 5 personas bastaría para dedicar el tiempo suficiente a cada indicador
- **Probar en empresas reales.** Debido a mis recursos me ha sido imposible probar la evaluación en empresas reales y después comparar con una puntuación oficial hecha a posteriori. Sería bueno realizar pre-evaluaciones en nuevas empresas con esta herramienta y ver de verdad su acierto.

11. Bibliografía

- I. EBC. Matriz del Bien común 4.1 (2014) [En línea]
<http://economia-del-bien-comun.org/sites/default/files/Matriz%20del%20bien%20com%C3%BAn%204.1.pdf>
- II. EBC. Matriz del Bien Comun 4.0 (2012) [En línea]
http://economia-del-bien-comun.org/sites/default/files/Matriz_Bien_Com%C3%BAn_4.0.pdf
- III. EBC. Manual Básico sobre el balance del Bien común (2012) [En línea]
<http://economia-del-bien-comun.org/sites/default/files/EBC-Manual%20basico%20sobre%20los%20criterios%204.0.5.pdf>
- IV. EBC. Manual de Diseño. Versión 1 en castellano (abril 2015)
- V. EBC. Ejemplos en español de informes del bien común en base a la versión 4.0. [En línea] <https://economydelbiencomun.org/ejemplos-de-balances/>
- VI. Stanford University. Machine learning [Curso en línea]
<https://www.coursera.org/learn/machine-learning>

[illegible]

[illegible]

Anexo II. Datos C1

Empresa	Horario de Trabajo		Configuración del Puesto		Salud Fisica y Seguridad	
Imanol Medi	Horario		Lugar	hogar	Prevision rie no	
85%	Fleixibilidad		Fen shui	no	Revisiones n no	
	Dias libres		Zonas Verde si		Comida trabajador	
	Vacaciones		Accesibilidac no		Formación s; no	
			Solicitudes si		Relajación no	
			Ergonomia si		Estudio Segu no	
			Espacios ampsi		Actividades \no	
			Ventilacion si		Seguro no	
	Salud Psiquica		Organización propia		Género	Población Desfavorecida
	Formación	facilitad:	Decisiones	horizontal	Remuneración	Alumnos
	Clima Bueno si		Programas d si		proporción	Desempleados
	Actividades (no		Herramientasi		Lenguaje Inclusivo	Inmigrantes
	Encuestas va no		Medidas Cor no		Talleres Formativos	movilidad reducida
			Encuestas sa no		Plan de inclusión	diversidad funcional
					Medidas Preventivas	exclusion
Empresa	Horario de Trabajo		Configuración del Puesto		Salud Fisica y Seguridad	
LMC	Horario		Lugar	hogar	Prevision rie no	
93%	Fleixibilidad		Fen shui	no	Revisiones n si	
	Dias libres		Zonas Verde no		Comida trabajador	
	Vacaciones		Accesibilidac no		Formación s; si	
			Solicitudes si		Relajación no	
			Ergonomia si		Estudio Segu no	
			Espacios ampsi		Actividades \si	
			Ventilacion si		Seguro no	
	Salud Psiquica		Organización propia		Género	Población Desfavorecida
	Formación libre		Decisiones	horizontales	Remuneración	Alumnos
	Clima Bueno si		Programas d si		proporción	Desempleados
	Actividades (si		Herramienta no		Lenguaje Inclusivo	Inmigrantes
	Encuestas va no		Medidas Cono		Talleres Formativos	movilidad reducida
			Encuestas sa no		Plan de inclusión	diversidad funcional
					Medidas Preventivas	exclusion

[illegible]

Empresa	Horario de Trabajo	Configuración del Puesto	Salud Física y Seguridad		
Oceanografía	Horario Democrático	Lugar oficina	Prevención riesgos	si	
	74% Flexibilidad personal	Fen shui	no	Revisión médica	si
	Días libres	Negociados	Zonas Verdes	no	Comida empresa
	Vacaciones	Adaptadas	Accesibilidad	no	Formación salud
			Solicitudes	si	si
			Ergonomía	si	Relajación
			Espacios amplios	si	Estudio Seguridad
			Ventilación	si	Actividades y talleres
				Seguro	si
				no	

	Salud Psíquica	Organización propia	Género	Población Desfavorecida
Formación	empresa	Decisiones jerárquicas	Remuneración	Igualitaria
Clima	Bueno	Programas de	proporción	50
Actividades		Herramientas	Csi	Lenguaje Inclusivo
Encuestas		Medidas Comunes	no	Talleres Formativos
		Encuestas	satis	Plan de inclusión
				Medidas Preventivas

Empresa	Horario de Trabajo	Configuración del Puesto			Salud Fisica y Seguridad			
Limonium	Horario	Establecido	Lugar	mixto	Prevision riesgos	si		
	76%	Flexibilidad personal	Fen shui	no	Revisiones medicas	si		
		Dias libres eleccion	Zonas Verdes	si	Comida	trabajador		
		Vacaciones adaptadas	Accesibilidad	no	Formación salud	si		
			Solicitudes	si	Relajación	no		
			Ergonomia	no	Estudio Seguridad	si		
			Espacios amplios	si	Actividades y taller	si		
			Ventilacion	si	Seguro	no		

	Salud Psiquica	Organización propia	Género	Población Desfav
	Formación empresa	Decisiones adaptado	Remuneración igualitario	Alumnos no
	Clima Bue si	Programas de c si	proporción 25	Desemple si
	Actividades si	Herramientas C si	Lenguaje Inclusivo no	Inmigrant no
	Encuestas si	Medidas Comu si	Talleres Formativos si	movilidad si
		Encuestas satis no	Plan de inclusión no	diversidad no
			Medidas Preventiv no	exclusion no

[illegible]

Empresa	Horario de Trabajo	Configuración del Puesto		Salud Física y Seguridad			
glugler	Horario	establecimiento	Lugar	mixto	Prevención	no	
?	Flexibilidad	personalización	Fenómeno	si	Revisión	si	
	Días libres	establecimiento	Zonas Ver	si	Comida	empresa	
	Vacaciones	establecimiento	Accesibilidad	si	Formación	si	
			Solicitud	no	Relajación	no	
			Ergonomía	si	Estudio Ses		
			Espacios	asi	Actividades	si	
			Ventilación	no	Seguro	no	
	Salud Psíquica	Organización propia	Género	Población Desfavorecida			
	Formación empresa	Decisiones	consensus	Remuneración	igualitaria	Alumnos	si
	Clima Bueno	Programas	no	proporción	<25	Desempleo	no
	Actividades	Herramientas	si	Lenguaje	Ino	Inmigrantes	no
	Encuestas valorativa	Medidas	Csi	Talleres	Fno	movilidad	si
			Encuestas	si	Plan de in	diversidad	si
					Medidas	Fno	exclusion
						no	

Empresa	Horario de Trabajo	Configuración del PLSalud Física y Seguridad						
Can cet	Horario	Estalecido	Lugar	oficina	Previsione	si		
80%	Fleixibilidad	ninguna	Fen shui	si	Revisione	si		
	Dias libres	personale	Zonas Ver	no	Comida	empresa		
	Vacacione	establecio	Accesibili	si	Formació	si		
			Solicitude	si	Relajación	no		
			Ergonomi:	si	Estudio Sesi			
			Espacios a	no	Actividad	no		
			Ventilacio	si	Seguro	si		
	Salud Psiquica		Organización propia	Género		Población Desfavore		
	Formació	empresa	Decisione	jerarquiza	Remuner:	igualitaria	Alumnos	si
	Clima Bue	si	Programa:	si	proporció	50	Desemple	si
	Actividade	si	Herramier	si	Lenguaje	si	Inmigrant	si
	Encuestas	si	Medidas	Csi	Talleres	Fsi	movilidad	si
			Encuestas	si	Plan de in	si	diversidad	si
					Medidas	Fsi	exclusion	no

Empresa	Horario de Trabajo	Configuración del PLSalud Física y Seguridad
Subte	Horario proyectos	Lugar trabajador
70%	Fleixibilidad familiar	Fen shui ?
	Dias libres personal	Zonas Ver ?
	Vacaciones ?	Accesibilid ?
		Solicitud ?
		Ergonomía ?
		Espacios a ?
		Ventilacio ?
e	Salud Psiquica	Organización propia
	Formació libre	Decisione horizontal
	Clima Bueno	Programa no
	Actividad no	Herramienta no
	Encuestas no	Medidas Csi
		Encuestas no
		Medidas Fno
		Medidas Fno

[illegible]

Anexo II. Datos C3

Empresa	Alimentacion		Movilidad		Cultura		Huella
Imanol M	Lugar	casa	Llegada	ninguno	Concienci	si	Huella
38	vegetarian	preocupar	Desplazan	privado	Ejemplari	si	<1.8
	cocina	si	Ubicacion	centro	Incentivos	no	
	comida	convencio	Contrataci	no	Material	no	
	Formacion	no	vestuario	no	Practicas	si	
	office	no	Reunione	no	Formacion	si	
	Proyecto	(no	Facilidad	no	Directivos	privado	
	restaurant	no	Facilidad	no	Analisis	no	
			Facilidad	/no			
			Apoyo fin	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
LMC	Lugar	casa	Llegada	ninguno	Concienci	si	Huella
19	vegetarian	preocupar	Desplazan	publico	Ejemplari	no	<1.8
	cocina	si	Ubicacion	centro	Incentivos	no	
	comida	local	Contrataci	no	Material	no	
	Formacion	no	vestuario	no	Practicas	no	
	office	no	Reunione	no	Formacion	si	
	Proyecto	(no	Facilidad	no	Directivos	privado	
	restaurant	no	Facilidad	no	Analisis	no	
			Facilidad	/no			
			Apoyo fin	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
Fundacion	Lugar	restaurant	Llegada	bicicleta	Concienci	si	Huella
19	vegetarian	ninguna	Desplazan	ninguno	Ejemplari	si	<5
	cocina	no	Ubicacion	centro	Incentivos	no	
	comida	convencio	Contrataci	si	Material	no	
	Formacion	no	vestuario	no	Practicas	no	
	office	no	Reunione	no	Formacion	si	
	Proyecto	(no	Facilidad	no	Directivos	?	
	restaurant	no	Facilidad	no	Analisis	no	
			Facilidad	/no			
			Apoyo fin	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
Mundo Ca	Lugar	oficina	Llegada	ninguno	Concienci	si	Huella
42	vegetarian	ninguna	Desplazan	ninguno	Ejemplari	si	ninguna
	cocina	no	Ubicacion	centro	Incentivos	no	
	comida	local	Contrataci	si	Material	no	
	Formacion	no	vestuario	no	Practicas	si	
	office	no	Reunione	no	Formacion	no	
	Proyecto	(no	Facilidad	no	Directivos	?	
	restaurant	no	Facilidad	no	Analisis	si	
			Facilidad	/no			
			Apoyo fin	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
AloeySalu	Lugar	casa	Llegada	ninguno	Concienci	si	Huella
55	vegetarian	preocupac	Desplazan	publico	Ejemplari	si	<4
	cocina	si	Ubicacion	centro	Incentivos	no	
	comida	local	Contrataci	si	Material	no	
	Formacion	no	vestuario	no	Practicas	si	
	office	no	Reunione	no	Formacion	si	
	Proyecto	(si	Facilidad	no	Directivos	publico	
	restaurant	no	Facilidad	si	Analisis	no	
			Facilidad	no			
			Apoyo fin	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
Eticus	Lugar	casa	Llegada	caminand	Concienci	si	Huella
42	vegetariano	ninguna	Desplazamientos	publico	Ejemplarid	si	desconoci
	cocina	no	Ubicacion	centro	Incentivos	no	
	comida	convencio	Contratacion	Cerc	Material	no	
	Formacion	no	vestuario	no	Practicas eco	si	
	office	no	Reuniones Virtua	no	Formacion	no	
	Proyecto Colab	no	Facilidad bici	si	Directivos	?	
	restaurant oficina	no	Facilidad Publico	si	Analisis	no	
			Facilidad Aparcam	no			
			Apoyo financiero	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
oceanogra	Lugar	oficina	Llegada	compartir	Concienci	si	Huella
38	vegetariano	ninguna	Desplazamientos	compartir	Ejemplarid	no	<5
	cocina	si	Ubicacion	afueras	Incentivos	si	
	comida	convencio	Contratacion	Cerc	Material	si	
	Formacion	no	vestuario	no	Practicas eco	si	
	office	si	Reuniones Virtua	no	Formacion	no	
	Proyecto Colab	no	Facilidad bici	no	Directivos	compartido	
	restaurant oficina	no	Facilidad Publico	si	Analisis	no	
			Facilidad Aparcam	si			
			Apoyo financiero	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
Limonium	Lugar	oficina	Llegada	publico	Concienci	si	Huella
75	vegetariano	mayoria	Desplazamientos	compartid	Ejemplarid	si	<1.8
	cocina	si	Ubicacion	?	Incentivos	si	
	comida	huerto	Contratacion	Cerc	Material	si	
	Formacion	si	vestuario	si	Practicas eco	si	
	office	si	Reuniones Virtua	no	Formacion	si	
	Proyecto Colab	no	Facilidad bici	no	Directivos	bicicleta	
	restaurant oficina	no	Facilidad Publico	no	Analisis	no	
			Facilidad Aparcam	si			
			Apoyo financiero	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
Market ea	Lugar	oficina	Llegada	publico	Concienciaci	si	Huella
63	vegetariano	mayoria	Desplazamientos	compartir	Ejemplaridad	no	<4
	cocina	no	Ubicacion	centro	Incentivos	si	
	comida	ecologica	Contratacion Cerc	si	Material	no	
	Formacion	si	vestuario	no	Practicas eco	si	
	office	si	Reuniones Virtua	si	Formacion	no	
	Proyecto Colab	no	Facilidad bici	no	Directivos	compartido	
	restaurante oficina	no	Facilidad Publico	si	Analisis	no	
			Facilidad Aparcar	no			
			Apoyo financiero	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
veritable	Lugar	oficina	Llegada	andando	Concienciaci	si	Huella
65	vegetariano	preocupac	Desplazamientos	publico	Ejemplaridad	no	<1.8
	cocina	no	Ubicacion	?	Incentivos	si	
	comida	ecologica	Contratacion Cerc	si	Material	si	
	Formacion	no	vestuario	no	Practicas eco	si	
	office	si	Reuniones Virtua	no	Formacion	no	
	Proyecto Colab	no	Facilidad bici	si	Directivos	andando	
	restaurante oficina	no	Facilidad Publico	si	Analisis	no	
			Facilidad Aparcar	no			
			Apoyo financiero	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
Delibera	Lugar	oficina	Llegada	Publico	Concienciaci	si	Huella
?	vegetariano	preocupaci	Desplazamientos	Publico	Ejemplaridad	si	<1.8
	cocina	no	Ubicacion	centro	Incentivos	?	
	comida	convencio	Contratacion Cerc	no	Material	?	
	Formacion	si	vestuario	no	Practicas eco	si	
	office	si	Reuniones Virtua	no	Formacion	?	
	Proyecto Colab	no	Facilidad bici	si	Directivos	?	
	restaurante oficina	no	Facilidad Publico	si	Analisis	?	
			Facilidad Aparcar	no			
			Apoyo financiero	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
waterscap	Lugar	oficina	Llegada	publico	Concienciaci	si	Huella
35	vegetariano	ninguna	Desplazamientos	publico	Ejemplaridad	si	ninguna
	cocina	si	Ubicacion	centro	Incentivos	no	
	comida	convencio	Contratacion Cerc	no	Material	no	
	Formacion	no	vestuario	no	Practicas eco	si	
	office	si	Reuniones Virtua	no	Formacion	no	
	Proyecto Colab	no	Facilidad bici	si	Directivos	publico	
	restaurante oficina	no	Facilidad Publico	si	Analisis	no	
			Facilidad Aparcar	no			
			Apoyo financiero	no			

Empresa	Alimentacion		Movilidad		Cultura		Huella
gluger	Lugar	oficina	Llegada	compartir	Conciencia	si	Huella
?	vegetarian	mayoria	Desplazam	compartir	Ejemplar	si	<1.8
	cocina	no	Ubicacion	centro	Incentivos	si	
	comida	ecologica	Contrataci	no	Material	si	
	Formacion	si	vestuario	no	Practicas	si	
	office	si	Reuniones	Virtuales	Formacion	si	
	Proyecto	(si	Facilidad	si	Directivos	?	
	restaurant	si	Facilidad	si	Analisis	no	
			Facilidad	/no			
			Apoyo fin	si			
Empresa	Alimentacion		Movilidad		Cultura		Huella
CanCet	Lugar	oficina	Llegada	privado	Conciencia	si	Huella
25	vegetarian	ninguna	Desplazam	privado	Ejemplar	no	<5
	cocina	si	Ubicacion	fuera	Incentivos	no	
	comida	convencio	Contrataci	no	Material	si	
	Formacion	no	vestuario	no	Practicas	si	
	office	si	Reuniones	no	Formacion	si	
	Proyecto	(no	Facilidad	no	Directivos	privado	
	restaurant	no	Facilidad	no	Analisis	no	
			Facilidad	/si			
			Apoyo fin	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
Subetica E	Lugar	oficina	Llegada	caminand	Conciencia	si	Huella
83	vegetarian	mayoria	Desplazam	bicicleta	Ejemplar	si	<5
	cocina	no	Ubicacion	centro	Incentivos	si	
	comida	huerto	Contrataci	si	Material	si	
	Formacion	si	vestuario	no	Practicas	si	
	office	si	Reuniones	no	Formacion	si	
	Proyecto	(si	Facilidad	si	Directivos	?	
	restaurant	no	Facilidad	si	Analisis	si	
			Facilidad	/no			
			Apoyo fin	no			
Empresa	Alimentacion		Movilidad		Cultura		Huella
Pauma	Lugar	oficina	Llegada	privado	Conciencia	si	Huella
11	vegetarian	ninguna	Desplazam	compartir	Ejemplar	no	ninguna
	cocina	no	Ubicacion	afueras	Incentivos	no	
	comida	convencio	Contrataci	no	Material	si	
	Formacion	no	vestuario	no	Practicas	si	
	office	si	Reuniones	no	Formacion	si	
	Proyecto	(no	Facilidad	no	Directivos	privado	
	restaurant	no	Facilidad	no	Analisis	si	
			Facilidad	/no			
			Apoyo fin	no			

0. English Version

Summary

This study is focus on new technologies application to the classifying process developed by the Economy theory of the Common Good. This economic theory developed in Austria has created a different classifying process of Companies to the ones currently established. In this process, ethical and ecological criteria are the priority versus the economical ones. Along years, an audit system has been developed by experts. Achieving a development economical model where Companies are interconnected and could benefit among them, providing a product to consumers that contribute to the Common Good is the main objective.

In this process, that nowadays is a manual process, I have identified an opportunity to use some technologies to *a priori* save time and advance prior work.

Assessment of each Company is performed under more or less standardized criteria, but is dependent on specific auditor perception. So that scores to Companies are established attending to 17 topics. Each of these topics is associated with different criteria that will have different weight in the valuation. At this moment, when the assessment of the audit Company is made I have identified the opportunity to apply a technology of automatic learning.

The main idea is to study the chance to automate these assessments through predictive algorithms.

When all documents reviewed and previously to the assessment, many indicators and attributes can be obtained to identify where to investigate to preview a final valuation. With all these attributes, data mining is performed to all data provided by the Companies. Besides this and the use of algorithms of automatic learning, a predictive model is searched, previously to the audit to be used as an initial draft and in the future with more data, might be uses even as a Company assessment system, and so many hours of analysis would be unneeded.

Results of this investigation will be useful for the Companies as the first steps with this type of application.

1. Introduction and objectives

1.1. Description

As explained in the title, the objective is to develop a system to classify companies attending to Common Good Economic theory. This work of study and development will be performed through data mining and automatic learning.

This theory implies an evaluation model of the company far from established rules and proposes a classification based on social good and sustainability. The theory provides tools to classify companies used by auditors to evaluate them in an adequate manner.

CGE organization has documents and trainings available for those interested in auditing companies. These persons perform social audits to companies that want to undergo them. Each indicator has rules and associated scales that provide an approximate score. This classification process is expensive and slow and there are not always enough people to audit, analyze companies and make the reports. An additional problem is that each evaluation process is individual and depends on various factors, one of them is the person performing the evaluation. Although they have rules to evaluate each indicator, these may be subject to interpretation and are not completely accurate.

The objective of this project is to investigate if a system of automatic evaluation that allows performing previous evaluations is possible to be developed. As previous evaluations and results are available, these are the ones considered to start the evaluation to obtain a predictive model close to reality.

When a theoretical model is applied, it is not so accurate from a practical perspective, thus, it is decided to proceed from examples and not from the theoretical model.

It is key to highlight the objective of this project is to investigate if taking account of a few indicators, the automatization of the process can be achieved. To be clarified that what it is not pursued is a final model with all indicators.

One of the main parts in this project is to study all the evaluation documents and try to model a group of data useful for the study and, onwards, may provide conclusions for the system. The study process of documents is

manual and consists on reading and extract all measures or initiatives reflected on and affected by the indicators.

Final objective of this project is to apply several machine learning techniques and logarithms to all data previously compiled to obtain an automatic system of classification. Until now, there is no project and nobody had investigate to apply these techniques to theory. Therefore, this project is a first step towards complete automatization way.

1.2. Personal objectives

This project is a personal challenge of great interest for me. Therefore, apart from the academic objectives I will explain my personal motivations to develop this project.

Common Good economy proposes an alternative way to current valuation applied to companies. Nowadays, companies are evaluated according to the value of their incomes and their profitability. In this market dynamic, stock holders are the ones getting benefits of these companies profitable activities. In this scenario, decisions are usually made without considering company employees or even against them or their interests. Furthermore, decisions are made against the planet and its long-term sustainability.

Raising this premise, I consider necessary another way to value and give advantages to companies. And it is within this concept where the Common Good theory has a key role.

Although it is only an intermediate and reformist solution of current productive system, a different point of view is provided. At the following chapter, a more detailed explanation will be provided but a summary of my motivations about this subject needed to be explained.

This alternative option clearly needs investment and dedication. With this project, my objective is to provide a grain of sand to this theory and provide of a classification system to automatize the theory. This way people will be able to easily access and use these classifications to consume in a conscious and sustainable way.

At technical level, this project also implies a challenge for me, and a learning process to leverage in my professional career and world of work.

- Deep dive in data mining techniques
- Deep dive in machine learning techniques

- Learning to deal with customers
- Management of a project in the field
- Learning management of a research project

1.3. Phases of development.

8. Study: starting point is performing a process of study and indagation.

First of all, an investigation and study of Common Good economy was required. I already knew this theory, but not in depth and I had to study and investigate its theory. During this process I contacted its organization in Madrid and I had several meetings with the people in this organization. Through these meetings I knew his work and I got a more precise idea of the theory they wield and with the documents and videos they have available on their web I got the knowledge needed.

On second hand, I wanted to refresh and deepen my knowledge about automatic learning to face the development of the process in the proper way.

I reviewed the subjects I had taken on this type of technology (automatic learning and artificial intelligence). To expand a little more, I took a *machine learning* course through *Coursera* platform. This process has taken the longest period together with the research performed.

9. Recognition of the problem. Once the needed knowledge to face the objective has been adquired, I focused to find out how I could apply this technology to the evaluation process. Reviewing all the documentation and reading the company documents, I noticed two things that affected these. First thing is the time that requires a complete evaluation of each company. In some cases, audits lasted 100 hours. The second problem results from the evaluation itself. Reviewing the audits it is observed that, although the criteria are pre-established, they are not completely accurate and are not applied in the same way in the companies. This phase did not take too long, because I became aware while investigating the issue in question.

10. Analysis. Once I recognized the problem, I compiled data through the documents provided by the EBC on the companies. Through an exhaustive reading, I was obtaining destinies, data or attributes that were (or not) in the audits and that, a priori, would have served for the valuation. This phase has been slow and laborious since I had to read carefully and draw conclusions from all the PDFs I had available. Step by step I was extracting enough information to go to the design phase of the experiments.

11. Design. This phase mainly consisted of elaborating a test plan and generating the model on which I was going to work. As for the tests, I designed the machine learning algorithms I was going to try and what kind of tests I was going to perform with them. In the model design I focused on properly defining the attributes I had collected in the previous phase and elaborating a model in inputs and outputs according to what I needed. This model would undergo modifications throughout the investigation, but it served as a prototype in the first instance.

12. Investigation. Once everything was defined, I investigated all data that I had available through the *weka* platform. I performed all tests I had programmed at the beginning and then adding some more that I found useful along the process. Through this research I obtained the relevant conclusions to develop the predictive system.

13. Development. This last part mainly consisted on developing a small application to show how the predictive equation works. This part is mainly visual and the importance lies in the research and in the search of a prediction formula.

1.4. Media

Physical equipment (Hardware)

- Desktop computer : AMD FX™ processor-8320 Eight cores 3.5 GHz, RAM 8GB, hard disk 1 T B and SSD disk 256 GB.
- Portable computer: Asus, Intel® processor Core™ i5-6300HQ Quad-Core (6M Cache, 2.3GHz hasta 3.2GHz), RAM 8 GB and Hard disk 1 TB

• Software

- Office 2016
- Weka 3.8
- Windows 10

2. Common Good economical theory.

2.1. Introduction.

” The Common Good economic theory is an economic project open to Companies and promoted by the Austrian economist Christian Felber that pretends to implement and to develop a truly sustainable economy as an alternative to financial markets where companies have to play”.

The Common Good theory proposes an alternative economic model to the current one.

This model is focus on sustainability, ecology, business ethics and good deal that differs from the standard concept of a company. Initiated by Christian Felber, he defines it as an alternative to market capitalism and to the planned economy. This theory implies a review of current economic model and enriched it with a long term perspective where sustainability and development of common good for our planet is guaranteed.

2.2. Common Good Balance.


Common Good Balance is the main tool of this theory. Through this balance, the role of the company around Common Good can be checked. This tool measures the impact of a company on Common Good. The objective is to convert this type of balance into the main mean to evaluate the company and that it becomes the value to compare and help consumers to choose, to the Governments to hire and other companies to choose others as partners. This balance would have the same reliability than current balance, but it would focus on other points considered decisive in the long term.

This Common Good Balance replaces the traditional and economic one based on incomes and profitability. In fact, it is much wider as it measures many more factors than the purely economic one. Main thesis is that economic development goes into the background and it is not considered as the final goal.

With the economic gains, a service or product that is useful, sustainable and ethical can be provided. Therefore, this balance comes to measure if the objectives to which all the companies should aspire if all were measured by this pattern are met.

At present, this balance consists of 17 indicators. As they explain, it is not a pricing process evaluation, but they seek to be understandable and easy to be applied. Each of these indicators has a value ranging from 20 to 90 points. To evaluate each of these indicators there are internal tables with levels that help to approximate the values.

2.3. Common Good Matrix

Valor ▸ Grupo de contacto ▾	Dignidad humana	Solidaridad	Sostenibilidad ecológica	Justicia social	Participación democrática y transparencia
A) Proveedores	 A1: Gestión ética de la oferta/suministros				90
B) Financiadores	B1: Gestión ética de finanzas				30
C) Empleados inclusive propietarios	C1: Calidad del puesto de trabajo e igualdad 90	C2: Reparto justo del volumen de trabajo 50	C3: Promoción del comportamiento ecológico de las personas empleadas 30	C4: Reparto justo de la renta 60	C5: Democracia interna y transparencia 90
D) Clientes / productos / servicios / otras empresas	D1: Relaciones éticas con los clientes 50	D2: Solidaridad con otras empresas 70	D3: Concepción ecológica de productos y servicios 90	D4: Concepción social de productos y servicios 30	D5: Aumento de los estándares sociales y ecológicos sectoriales 30
E) Ámbito social: región, soberanía, generaciones futuras, personas y naturaleza mundial	E1: Efecto social/ Significado del producto/ servicio 90	E2: Aportación a la comunidad 40	E3: Reducción de efectos ecológicos 70	E4: Orientación de los beneficios al bien común 60	E5: Transparencia social y participación en la toma de decisiones 30
Criterios – Negativos	Quebrantamiento de las Normas de trabajo OIT -200	Compra hostil -200	Gran impacto medioambiental a ecosistemas -200	Evasión de impuestos -200	No revelación de todas las participaciones -100

3. Study proposal

3.1. Problem Statement

Several problems can be observed that are visible and allow to look for improvement options. The following are the problems I have seen during my research and previous work:

- **The process of creating balance sheet development becomes slow and cumbersome.** It requires a detailed analysis of all aspects of the company and besides it is necessary to take into account numerous factors for each valuation. This sometimes lasts 100 hours to get a first assessment. From the first balance the rest last less. There are companies that have that capacity, but there are others that can not always invest so much time in doing the study. This temporary problem also applies to the auditors. They have to spend much of his time studying companies. In addition, although the CGE is state-owned, it does not have subsidiaries in all regions of Spain. Therefore, there are areas where there will not always be people capable of carrying the process.
- **The valuation process tends to be subjective and suffers variations in the different companies.** In the internal tables, criteria and associated levels are not defined extremely precisely. Moreover, certain levels are too subjective and lend themselves to personal appraisals. In this way, valuations are produced that

differ widely between companies, even if they have similar characteristics and should have the same score. It is not completely clear and lends itself to the interpretation of the person who evaluates.

- **An specific training is required to perform the balance.** To carry out the balance, a specific training is required that allows to evaluate business context and perform the evaluation accordingly. It is true that this training is needed to produce a detailed report, but it might not be necessary to carry out an initial assessment that serves as a draft.

Definitively, the problems I consider associated to this type of assessment are **TIME, SUBJECTIVITY AND TRAINING**

3.2. Solution proposal

The idea is to work on a predictive model of evaluations. A series of equations or decision trees that serve to score companies, at least in an initial moment, without the need of a long process.

Since evaluation documents are already available, the idea is to work on them. The evaluation system is subjective but previous audits are available. Therefore, the proposal is to analyze each document and try to study all measures it contains for each given attribute and see what score they have obtained. This analysis must be manual, since it is necessary to infer data, not always written. It could be tried to be performed with some language processor, but information would be lost.

Once documents have been analyzed, a set of data will be available to serve as training. By having training data it is possible to work with them through processes of data mining and automatic learning. The objective of this study is to find the model from these training data that will serve us for future cases. Through different algorithms of Machine Learning will try to find the one that best model that suits the reality.

How does this solution affect the problem? It is simple. If you have equations to obtain *a priori* valuations, many companies will be able to make the jump to these types of balances without much effort. Time required to study the company and detail each of the measures is exponentially reduced by having an equation with some general attributes that serve for a first approach. In later studies when more balances are available, the problem of automating the entire process can be addressed. The subjectivity in the application is reduced since from the examples we obtain a model in which only one has to fill fields to obtain the score. And given the simplicity of this type of tool does not previous training

is required to obtain the balance of a company and start participating in the network of companies for the common good.

4. Investigation

The decision was made to model only some indicators of the matrix, and this was decided to avoid the great effort that had required studying the 17 attributes manually. The three Indicators chosen were the A1, C1 and C3. The A1 was chosen to be one of the most extensive, since a number of factors affected it. The C3 was chosen as it has many variables. And the C3 for being simple.

The process of data extraction is slow and meticulous. The documents containing the audits of the companies that had been carried out to date were read. Since only a few indicators had been selected, the reading was focused only on those. Through in-depth study and comparison, initiatives, actions or procedures that at first seemed to affect the final valuations were obtained. Initiatives encountered in audits were intersected with others to develop a complete data set.

At the same time, its relationship or not with the levels that a priori contributes the documentation that teaches how to value companies was studied.

With all this a list was built and endowed with attributes with defined options. Thus, a set of hypotheses that would determine the score of the indicator was prepared. This model is the one used to investigate and from which the prediction equations have been obtained.

The testing process has been methodical. The algorithms were tested with the same parameters for the different sets of data that were generated from the initial ones. For those who showed better results, more exhaustive tests were performed, but for the others a few were enough. Many times when the best algorithm was found, rapid tests with different modifications were performed to find the best data set. This allows finding better datasets.

Many times when the best algorithm was found, rapid tests with different modifications were performed to find the best data set. This allows finding better datasets.

5.1 Results

Ecuación A1

1.1753 * tamaño=reducido, grande, medio +
0.8275 * tamaño=grande, medio +
0.7888 * fijo=ecológico, local, ético, preocupación +
0.7888 * fijo=local, ético, preocupación +
0.7888 * fijo=ético, preocupación +

3.1558 * fijo=preocupación +
0.6885 * agua=ético, local, preocupación, ecológico +
0.6885 * agua=local, preocupación, ecológico +
0.6885 * agua=preocupación, ecológico +
0.7001 * agua=ecológico +
0.5109 * electricidad=preocupación +
1.598 * gestoría=ético, ecológico, preocupación, convencional +
1.598 * gestoría=ecológico, preocupación, convencional +
1.598 * gestoría=preocupación, convencional +
1.598 * gestoría=convencional +
1.2167 * prov_servicios=preocupación, ético, local, ecológico +
1.2167 * prov_servicios=ético, local, ecológico +
1.2167 * prov_servicios=local, ecológico +
1.0241 * prov_servicios=ecológico +
1.4966 * prov_material=preocupación, ético, local, ecológico +
1.4966 * prov_material=ético, local, ecológico +
1.4966 * prov_material=local, ecológico +
1.1436 * prov_material=ecológico +
3.0277 * seguros=ético, ecologico, preocupación, local +
2.4073 * seguros=ecologico, preocupación, local +
2.4073 * seguros=preocupación, local +
2.9846 * seguros=local +
1.9487 * transporte=publico, ético, local, ecológico, convencional +
1.3655 * transporte=ético, local, ecológico, convencional +
1.3655 * transporte=local, ecológico, convencional +
2.0803 * transporte=ecológico, convencional +
2.4368 * transporte=convencional +
1.2763 * meterial_oficina=convencional, local, ético, ecológico +
1.0716 * meterial_oficina=local, ético, ecológico +
1.0716 * meterial_oficina=ético, ecológico +
1.0716 * meterial_oficina=ecológico +
2.2233 * material=local, ético, convencional, ecológico +
2.4589 * material=ético, convencional, ecológico +
2.4589 * material=convencional, ecológico +
2.0158 * material=ecológico +
1.9273 * ahorro_ecológico=si +
0.8275 * eficiencia_energética=si +
1.4093 * reciclaje=si +
0.9264 * proteccion_laboral=si +
1.5072 * rechazo_ético=si +
0.1268 * formación=si +
0.9685 * estudio_impacto=si +
0.9921 * encuestas_clientes=si +
2.0049 * concursos=no +
1.5072 * precios_ajustados=si +
16.4852

Ecuación C3	
0.6557	* tamaño=autónomos, medio, reducido +
0.2922	* tamaño=medio, reducido +
3.7414	* tamaño=reducido +
5.1537	* lugar=casa, oficina +
0.3454	* lugar=oficina +
3.4501	* vegetariano=todo, preocupación, mayoría +
3.4501	* vegetariano=preocupación, mayoría +
3.6774	* vegetariano=mayoría +
1.9425	* comida=local, ecológica, huerto +
3.6289	* comida=ecológica, huerto +
3.883	* comida=huerto +
3.6774	* formacion_aliment=si +
7.4106	* colaborativo=si +
0.6557	* llegada=bicicleta, compartido, ninguna, empresa, público, caminando +
2.5071	* llegada=compartido, ninguna, empresa, público, caminando +
2.7944	* llegada=ninguna, empresa, público, caminando +
0.3512	* llegada=empresa, público, caminando +
0.3512	* llegada=publico, caminando +
3.9141	* llegada=caminando +
0.9368	* desplazamiento=privado, público, empresa, caminando, compartido, bicicleta +
-5.6625	* desplazamiento=público, empresa, caminando, compartido, bicicleta +
-1.2005	* desplazamiento=empresa, caminando, compartido, bicicleta +
-1.2005	* desplazamiento=caminando, compartido, bicicleta +
-1.2005	* desplazamiento=compartido, bicicleta +
5.3951	* desplazamiento=bicicleta +
1.0497	* ubicación=centro +
5.0867	* proximidad=si +
1.7735	* vestuario=si +
2.1662	* virtuales=si +
2.02	* fpublico=si +
3.776	* faparcamiento=si +
4.063	* ejemplaridad=si +
2.5802	* incentivos=si +
1.5045	* material=si +
11.2169	* practicas=si +
2.0676	* directivos=empresa, ninguna, público, compartido, caminando, bicicleta +
2.0676	* directivos=ninguna, público, compartido, caminando, bicicleta +
2.0676	* directivos=público, compartido, caminando, bicicleta +
2.2637	* directivos=compartido, caminando, bicicleta +
2.7634	* directivos=caminando, bicicleta +
2.4267	* directivos=bicicleta +
-0.96	* análisis=si +
2.0668	

Results from C1 attribute are not very encouraging. As already explained, C1 was chosen since it was the longest attribute in the table. The idea was to test if it could be adapted to such complicated attributes. In A1 it has been achieved, but with C1 it has not been possible.

6. Conclusions

If I had to give a conclusion to define everything I would say that the work has been enriching and fruitful but there is much to do and a long way to go ahead. In fact, I have decided to add an extra point to comment on the next steps that are needed in this regard in case someone decides to continue the work I have started.

I believe that applying automatic learning to the evaluation process is a good idea in the future to be able to solve the problems that I described at the beginning of the document. The equations reduce the time needed to evaluate, eliminate the need for previous training for the completion of documents and do not require people visiting regularly to companies.

It is a technology that fits to this type of initiatives in process of growth and constant creation. It is very useful because the documents produced can generate equations that fit the reality and that can be changed in case of changing criteria and or valuations. In case some criterion lost weight, it would enough to re-adjust data and re-run the logarithms on the set. Or even to modify the equations by lowering the specific weight to the given variable would also be worth. Taking into account that some parameters are changing from time to time this adaptability is important.

It is still an initiation and the results are neither definitive nor totally conclusive. As can be seen in the error of the equations still lack precision in the prediction, the errors are high. This I think is due to several circumstances. I will comment on two problems first that refer to the theoretical part or the elaboration of system.

The first is human error. Part of my work has consisted in systematizing, in the form of variables, some documents based on the CGE manuals and the comparison between them. This process is subject to errors. I may have interpreted variables that do not influence or I may have created attributes that had no consistency that simply did not fit into reality.

Secondly, I think there is no consistency or scalability in the documents. That is, I believe that some companies are overvalued and others are undervalued. During the study of the variables I perceived that companies that *a priori* does not comply excessively with the values or that does not have measures at a particular point has inflated scores. On the other hand, others seem to have the same score with many better initiatives. It would be good to review or at least ensure that the general criterion

stated in the manuals is met. If the evaluation method were more precise it would be possible to obtain more realistic scores, which is mainly what I was looking for with this project and that at certain manner may work.

Going on with the issues I have found in carrying out the work, I will highlight two problems that are somehow more technical.

First one is the lack of training examples. In total, I have had 16 documents from 16 different companies, all of quite different sizes. Besides, 2 of them did not have the valuation in number form. That is, there were 14 examples. This has been one of the main problems in reaching conclusions. With such a small set it is difficult to create an equation in which we can ensure that the score is accurately obtained for any case. Second and related to the first one appears the problem of overfitting.

Having so few examples of training I can get an equation that fits the previous examples but that it were not valid for new companies. But although problems are relatively important have an easy solution, as time goes on and more documents were available, more detailed testing can be done with several training sets and get better equations.

On a personal perspective, I think that this work has enriched me a lot. I have learned to work on the ground. Having meetings with the organization, communicating them advances, requesting documentation or even consulting doubts while doing the work has been a very good experience. A very real experience like when working with a client in a computer project. In addition, it has been very enriching to be able to apply or investigate the application of a technology in a field where it has not been applied yet.

It has served me to learn, both about the Common Good economy and on automatic learning. During the work I was studying an online Course of Machine learning that was very useful to focus much more on data management than to artificial intelligence as I have learned in my degree. In addition, the fact of being able to apply it to a real case has been very satisfactory since I have been able to see its possibilities and its real usefulness. In addition, I have had to work hard on CGE and read and study their materials to understand them and then apply them to my development.